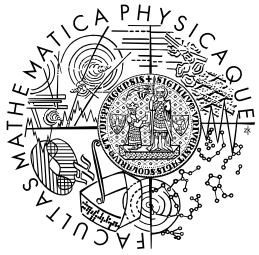


FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

MASTER THESIS

Gabriela Martinovicová

Study of fake-tau background with the ATLAS experiment

Institute of Particle and Nuclear Physics

Supervisor of the master thesis: Mgr. Vojtěch Pleskot, Ph.D.
Study programme: Physics
Study branch: Nuclear and Subnuclear Physics

Prague 2021

I declare that I carried out this master thesis independently, and only with the cited sources, literature and other professional sources. It has not been used to obtain another or the same degree.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In ..... date ....

Author's signature

First of all, I have to say a huge thank you to my supervisor Mgr. Vojtěch Pleskot, PhD. for hours spent consulting this thesis, all his help, expertise, and mainly the active approach in solving problems that arose during the work on this thesis. I would also like to thank him for his support and guidance throughout my studies.

Great gratitude goes to my parents and my family. Everything would be much harder without their endless support.

Title: Study of fake-tau background with the ATLAS experiment

Author: Gabriela Martinovicová

Institute: Institute of Particle and Nuclear Physics

Supervisor: Mgr. Vojtěch Pleskot, Ph.D., Institute of Particle and Nuclear Physics

Abstract:

The τ-leptons are the important final-state components, not only in the Standard Model processes but also in the processes beyond the Standard Model studied at the ATLAS experiment at CERN. They are characterized by mostly decaying into hadrons with one or three charged particles and, in most cases, with at least one neutral pion in the final-state. Due to their short decay length, only their decay products are observed in the detector. Jets naturally fake hadronically decaying τ leptons, so it is necessary to estimate such a fake-τ background. The Fake Factor method uses a correction factor, called fake factor (FF), measured from the data and applied to the data to estimate the fake-τ background in a given signal region. One of the complications is that FF differs for τ candidates faked by jets derived from quarks or gluons and thus must be measured in the control region with the same fraction of quark jets as in the signal region. The solution to this problem is the universal FF method, which proposed that from the FFs measured in samples with a large difference in the fraction of jets originating from quark (Z+jets) and gluon (multi-jet), the FF in any particular signal region can be interpolated. This thesis focuses on measuring FF in Z+jets events, in which most of the fake-τ candidates originate from quarks, and studying various aspects that may affect this measurement.

Keywords: CERN, ATLAS, LHC, tau lepton, fake-tau

Contents

Introduction

The processes with the τ leptons in the final state are essential for many Standard Model measurements as well as for searches for physics beyond the Standard Model. Due to its short decay length, τ leptons produced in proton-proton (pp) collisions within the ATLAS experiment at the LHC, can be observed only in the form of their decay products. The τ lepton decays either leptonically or more often hadronically. In analyses using hadronically decaying τs, jets misidentified as τ leptons are crucial source of background. The Monte Carlo is not able to properly model the fake-τ background. Thus, data-driven methods are preferred. In this master thesis, we focus on one of the main methods used in the ATLAS experiment: the fake factor method.

The fake factor method estimates the fake background by measuring the correction factors to be applied to data, where correction factors are referred to as fake factors. However, due to the different properties of quark- and gluon-initiated jets, the fake factor measured in samples rich in quark- or gluon-initiated jets are also different. Hence, the fake factor should be measured in a control region with a similar fraction of quark- and gluon-initiated jets to the signal region, which is quite challenging to accomplish. The solution to this problem is the universal fake factor method introduced in this thesis. This method proposes that the fake factor for any particular signal region can be interpolated knowing the quark fraction of a given signal region from the fake factors measured in samples with a significant difference in quark and gluon fraction. A suitable sample dominant in the production of gluon jets is the multi-jet sample, and for quark-initiated jets, it is the Z+jet sample.

This master thesis focuses on the measuring the fake factor in $Z(\mu\mu)$+jet events in data collected in Run 2 (2015-2018) with center-of-mass energy of proton-proton collisions $\sqrt{s} = 13$ TeV. Before the final measurement of the fake factor, many attributes had to be considered and studied, such as dependency study on p_{T} or τ decay modes, or the number of close-by objects that could affect leading τ-candidate and others. All of these studies are described in this thesis.

1. Theoretical Background

1.1 Standard Model of Particle Physics

It seems that everything in our universe is made from few building blocks known as fundamental particles and interactions that act upon them that do not appear to be reducible to more basic interactions. Classification of fundamental particles and three of these fundamental interactions can be described by the Standard Model (SM) of particle physics. In the following lines, only a brief introduction to the SM theory will be provided, mainly based on [1], [2]. The formulation of SM began with the unification of electromagnetic and weak interactions between elementary particles by Glashow-Weinberg-Salam theory of electroweak interactions, which was later extended by the quantum chromodynamics (QCD), a theory of strong interaction.

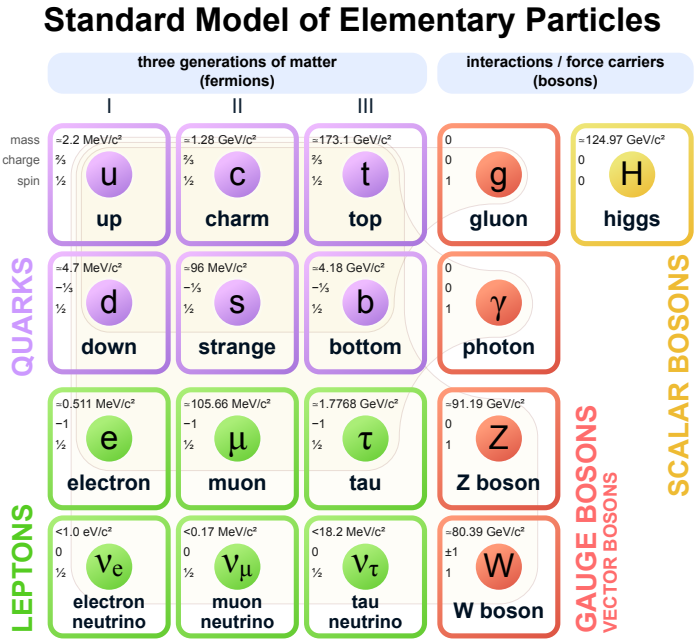


Figure 1.1: Standard Model of particle physics [3]

The SM provides a classification of elementary particles (see Fig. 1.1) and describes the known behavior of these particles. All of the matter is made of 12 elementary particles with half-integer spin, which follow Fermi-Dirac statistics, known as fermions, respecting the Pauli exclusion principle. These are further divided by whether they carry color charge and consequently interact via the strong interaction or not. Fermion carriers of the color charge are called quarks, and the remaining fermions are known as leptons. The SM theory proposes three generations of leptons: electron e – electron neutrino ν_e, muon μ – muon neutrino ν_μ, tau τ – tau neutrino ν_τ. Similar to the leptons, there are three generations of quarks: up – down, charm – strange, top – bottom (beauty). Only the

first generation of fermions forms stable matter. For every fermion, there is an antiparticle with the same mass but opposite identifying quantum numbers.

Particles with spin 1, which obey Bose-Einstein statistics, known as gauge bosons, provide for interactions between fermions. As mentioned at the beginning of this section, there are four known fundamental interactions in nature: gravity, electromagnetic, weak, and strong interaction. Only three of these forces are described by SM. Mathematically, the SM is $\mathrm{SU(3)_{color} \times SU(2)_L \times U(1)_Y}$ renormalizable gauge theory effected by the Higgs mechanism. The Higgs mechanism generates the masses of all massive elementary particles, including vector bosons $W^{\pm}, Z$ when they interact with the Higgs field. The scalar Higgs boson is an associated particle with the Higgs field. The spontaneously broken $\mathrm{SU(2)_L \times U(1)_Y}$ gauge group describes electroweak interactions, where electromagnetic interaction is intermediated by massless photon γ and weak interaction is carried by $W^{\pm}$ and Z^0. The strong interaction is described by the SU(3) group with force carrier gluon g.

As already mentioned, quarks carry a color charge, more precisely one of 3 colors/anticolors. Color is the main quantum number of the strong force, and the color state of quarks can be changed by either absorbing or emitting gluon. Gluon carries both colors and anticolors simultaneously. Hence, it can be found in one of the eight possible color states. Quarks create colorless bound states called hadrons, and in nature, quarks occur only in this form. Moreover, hadrons can be divided into two main groups based on a number of quarks they consist of: mesons consisting of $q\bar{q}$ and baryons consisting of qqq or $\bar{q}\bar{q}\bar{q}$. There are also other more rare hadrons as tetraquarks or pentaquarks.

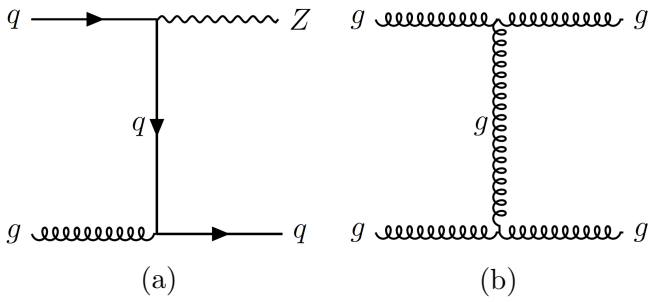


Figure 1.2: Feynman diagrams for production of (a) Z+ jets and (b) dijet at a tree-level

This thesis focuses on the estimation of fake background by jets misidentified as τ leptons. τ lepton is the heaviest among the leptons, $m_\tau \doteq 1.776$ GeV [4] and the most unstable with decay length of $c\tau = 87.03\mu$m. It also differs from other leptons due to its hadronic decay modes (approx. 65% decay modes are hadronic). The estimation of fake background is carried out by calculating scale factors - fake factors, which are dependant on the fraction of quark- or gluon-initiated jets. For fake factor measurements, Z+jets and multi-jets events were used as there is a large difference in the fraction of jet origin for these events. In Fig. 1.2 reader can see corresponding Feynman diagram for production of given events at a tree-level, where the Feynman diagram for dijet production is

only a representative of the greater number of tree-level diagrams. Z boson with mass $m_Z \doteq 91.19$ GeV decay dominantly into hadrons, but also into leptons with branching ratio about 3% for each massive l^+l^- pair.

1.2 Jets and anti-k_t algorithm

In QCD, quarks and gluons are identified with R. Feynman's partons. Interactions between partons are described by the hard scattering process, with partons also in the final state. Due to color confinement which only allows colorless states, parton cannot be observed experimentally. After hard scattering, fragments carrying some of the color charge have to form colorless hadrons - this process is called hadronization. The observable that provides the connection between the partons resulting from hard scattering and detected hadrons is called jet.

For jet identification, jet algorithms are used. It maps the momenta of the final state hadrons into the momenta of the jets. There are two main types of jet algorithms: sequential clustering algorithms and cone algorithms. The ATLAS Collaboration uses the anti-k_t jet clustering algorithm [5]. The anti-k_t algorithm is one of the sequential recombination jet algorithms that are infrared and collinear (IRC) safe. The IRC safety means that the given algorithm is insensitive to the emission of a soft gluon and also to the splitting of a hard particle.

The sequential clustering algorithms are based on finding the pair of partons with the smallest distance and recombine these two partons into one. For input, they use the final state collinear particles in a shower. This process is repeated until a stopping criterion is reached. In all sequential clustering algorithms, two distance measures are defined. One, distance d_{ij} between partons i and j and second, distance d_{iB} between parton i and the beam (B):

$$d_{ij} = \min(p_{\mathrm{T}i}^{2p}, p_{\mathrm{T}j}^{2p}) \frac{R_{ij}^2}{R^2}, \tag{1.1}$$

$$d_{iB} = p_{\mathrm{T}i}^{2p}, \tag{1.2}$$

where $R_{ij} = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}$ is the distance of partons i and j in $\eta \times \phi$ plane and $p_{\mathrm{T}i}, \eta_i$, and ϕ_i are the transverse momentum, rapidity, and azimuth angle of particle i, respectively. Radius parameter R relates with the final size of the jet in $\eta \times \phi$ plane, and parameter p corresponds to the relative power of the energy versus geometrical scales. After computing distance measures for each parton pair, the clustering continues with the identification of the smallest among the distances d_{ij} and d_{iB}. If the smallest number is d_{ij}, the partons i and j are combined into one. In the case of d_{iB} being the smallest one, i is called jet and removed from the list of entries. The whole procedure is repeated until no entries are left.

The anti-k_t algorithm takes $p = -1$. It can be seen from Eq. 1.1 that the anti-k_t algorithm prefers clustering a soft particle with a high-p_T (hard) particle long before soft particle clusters with another soft particle. If there are no two hard

particles within a distance of $2R$, the hard particle clusters all soft particles within the radius R. In this case, the anti-k_t algorithm provides a perfectly conical jet. On the other hand, if there are two hard particles and their distance in $\eta \times \phi$ plane is within $(R, 2R)$, jet with higher p_T will be fully conical, and the other one will be partly conical because of overlapping with the first jet and if jets have equal p_T the overlapping part will be divided equally between them. For $R_{12} < R$, hard particles 1 and 2 will result in a single jet with cone centered on p_T of particle with higher p_T or united cones of each hard particle plus a final jet cone.

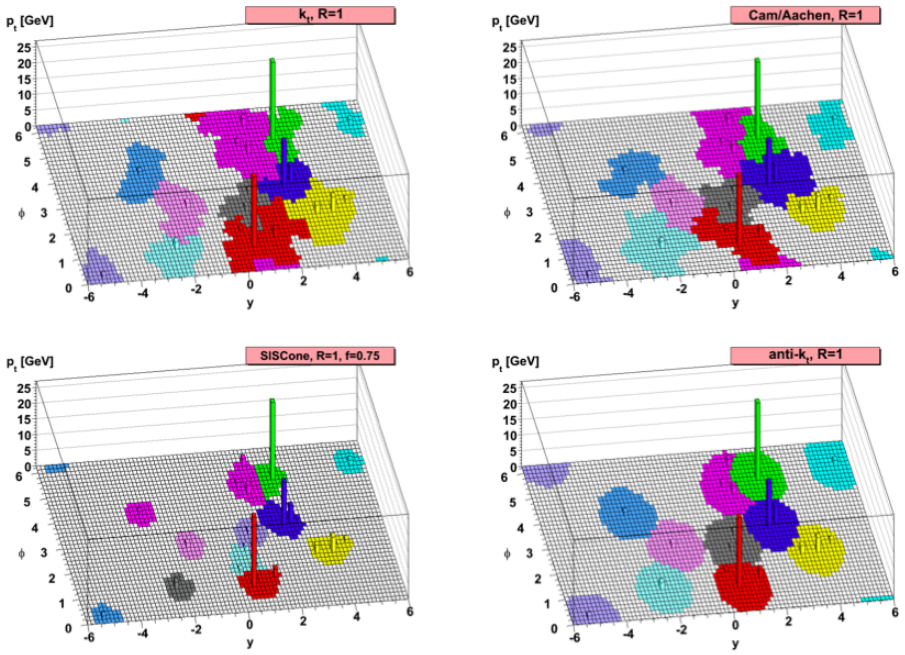


Figure 1.3: Comparison of the jet reconstruction algorithms. SISCone is one of the cone algorithms. Others are sequential clustering algorithms: $k_t(p = 1)$, Cam/Aachen $(p = 0)$ and anti-k_t $(p = -1)$ [5]

2. The ATLAS Experiment

2.1 CERN

European Organization for Nuclear Research (CERN - Conseil Européen pour la Recherche Nucléaire) is one of or even the most important organization in the field of research of elementary particles and structure of matter. Based in Geneva on the Franco-Swiss border, CERN was established in 1954 by 12 founding Member States, currently having 23 member states.

CERN is best known for its discovery of the Higgs boson, but also carriers of the weak force - Z and W bosons were discovered here and many others. CERN hosts several experiments, and one of the major four is the ATLAS experiment. However, CERN is not only known for its particle accelerator and discoveries of multiple particles, but also the World Wide Web was invented in the CERN laboratory.

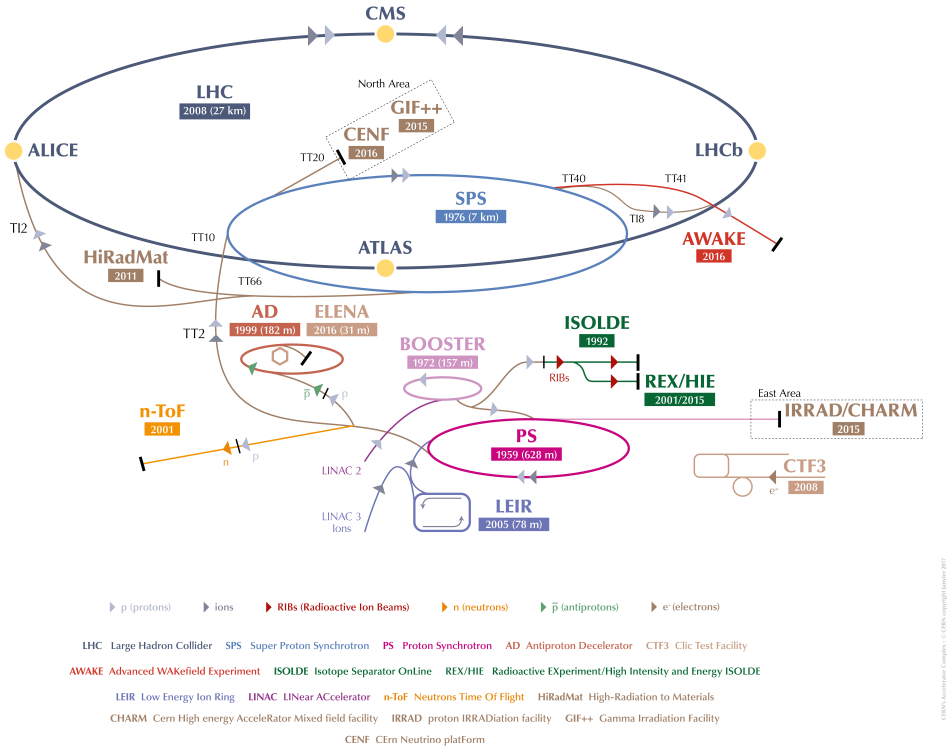


Figure 2.1: Cern's accelerator complex [6]

2.2 Large Hadron Collider

CERN's accelerator complex is a succession of many accelerators, and its nucleus is the Large Hadron Collider (LHC).

In this accelerator, which is located inside of the almost 27-kilometer long circular tunnel, are primary accelerated proton beams. These proton beams are formed by so-called bunches, more precisely, designed value for LHC was 2808 bunches and each bunch containing 1.2×10^{11} protons at start [7]. The source of protons is a simple hydrogen gas bottle, which is stripped of its electrons by the electric field. The proton beams are gradually accelerated in the chain of accelerators (see Fig.2.1), starting in Linac2 and then injected into Proton Synchrotron Booster (PBS) followed by Proton Synchrotron (PS) and Super Proton Synchrotron (SPS). In SPS, the proton beam is accelerated to 450 GeV and finally transferred to LHC.

In LHC, proton beams circulate in two separate beam pipes, clockwise in one and anticlockwise in the other. These tubes are kept at ultrahigh vacuum (pressure is in the order of 10^{-10} to 10^{-11} mbar [8]), and a strong magnetic field (the main dipoles generate the magnetic field of 8.3 T) maintained by 9593 superconducting electromagnets is used for guiding and focusing the beams [9]. The electromagnets are built from coils of niobium-titanium (NbTi) cables that operate in the superconducting state. Therefore they conduct electricity without resistance or loss of energy, and for NbTi to become superconducting, the magnets need to be chilled to 1.9 K (-271.3°C).

For LHC Run 2, which took place in 2015-1018, center-of-mass energy of *pp* collisions was $\sqrt{s} = 13$ TeV, and that means that proton beams in each beam pipe are accelerated up to the energy of 6.5 TeV. For protons, it takes 20 minutes to reach this maximum energy. After reaching this energy, the beams circulate in the LHC ring for many hours and make 11 245 turns around the accelerator's circumference per second. Not only protons are accelerated in LHC but also heavy ions, which have designed energy of 2.8 TeV for nucleon.

The designed LHC instantaneous luminosity for protons is $10^{34}\text{cm}^{-2}\text{s}^{-1}$ and due to the high luminosity, a multiple proton-proton collisions may occur in single bunch crossing (pile-up). With designed LHC collision frequency 40 MHz (bunch crossing every 25 ns) and pile-up, up to two billion proton-proton collisions occur.

The LHC operates four simultaneous detectors distributed around its perimeter - ATLAS, CMS, ALICE, and LHCb. In these places, the beam tubes' penetration occurs, which allows the collisions of accelerated particles.

2.3 The ATLAS detector

The ATLAS detector is a general-purpose detector at LHC. The acronym ATLAS stands for A Toroidal LHC Apparatus. ATLAS is the biggest detector operating on the accelerator that was built till now. Its dimensions are 25 m in height and 46 m in length. With respect to the interaction point, the ATLAS detector has nominal forward-backward symmetry and also symmetry in the azimuthal angle.

Since not every particle born in *pp* collisions interacts with a matter in the same way, the ATLAS detector has to consist of different subdetectors (subsystems) (see Fig.2.2). There are four main subsystems: inner detector, calorimetry, muon

spectrometer. Combined data volume from observed proton-proton collisions corresponds to more than 60 terabytes. It is not possible yet to store such big amounts of data, therefore after the detection, the ATLAS trigger system preselects events and reduce the data to manageable levels.

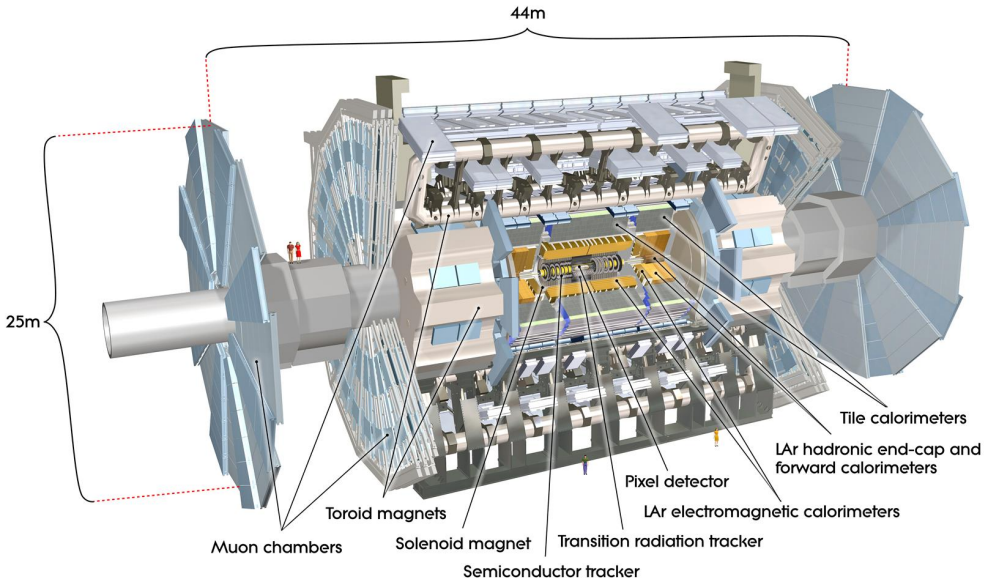


Figure 2.2: ATLAS detector layout [10]

2.3.1 ATLAS Coordinate System

The coordinate system of the ATLAS detector is a right-handed Cartesian system. The interaction point is taken as the origin of the coordinate system. The z-axis is parallel to the beam pipes. The $x-y$ plane is perpendicular to the beam direction and is referred to as the transverse plane. The positive x-axis points from the interaction point to the center of the LHC, and the positive y-axis points from the interaction point toward the Earth's surface. The momentum of the particles measured in this plane is called the transverse momentum p_{T}.

The transverse plane is often given in spherical coordinates θ, ϕ. The azimuth angle ϕ is measured with respect to the positive x-axis in the transverse plane. The polar angle θ is defined as the angle from the positive z-axis. The pseudorapidity is defined as $\eta = -\ln\tan(\theta/2)$. The distance ΔR is defined in the $\eta - \phi$ plane as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ [10].

2.3.2 Inner Detector

To achieve the momentum and vertex resolution requirements, despite huge track density, which is due to a significant number of particles (approx. 1000 particles every 25 ns) emerging from the collision point within $|\eta| < 2.5$, high-precision

measurements with fine detector granularity are needed. The Inner Detector (ID) with its primary components silicon Pixel Detector (PD), Semiconductor Tracker (SCT), and Transition Radiation Tracker (TRT) is used to achieve this.

ID is located in a 2 T solenoidal magnetic field (generated by the central solenoid, which is 5.3 m long with a diameter of 2.5 m), which causes the charged particle to deflect, and from its path, it is possible to determine the electric charge and calculate the momentum. Since neutral particles do not ionize the environment they pass through, it is impossible to detect them in this detector.

Each of the detector components is formed by two parts, a part which is parallel to the beam axis, called the barrel, and a part perpendicular to the beam axis, called the end-caps. The barrel part consists of concentric cylinders, and end-caps are formed by disks. The layout of the ID is illustrated in figure 2.3.

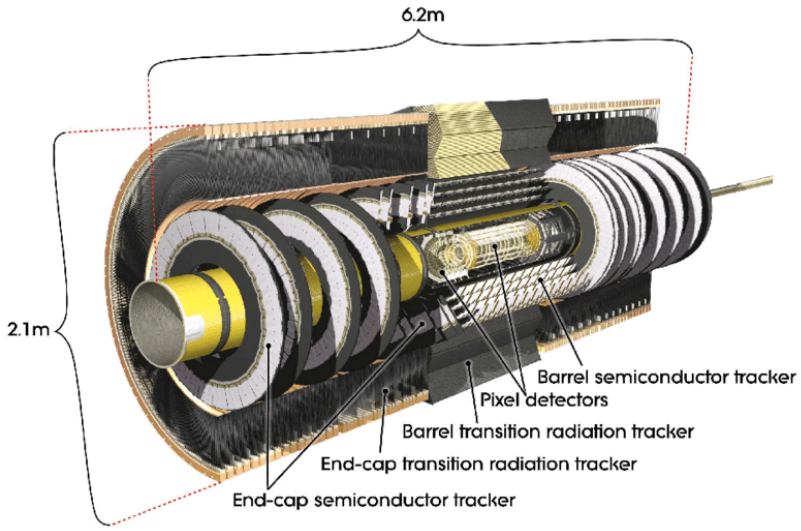


Figure 2.3: Cut-away view of the ATLAS inner detector [11]

PD and SCT are precision tracking detectors with coverage of the pseudorapidity range $|\eta| < 2.5$. PD is located in the center of ATLAS. It provides measurements with high accuracy even at the smallest possible distance from the vertex. It allows the detection of particles with a short life, such as B-hadrons. It consists of 80 million pixels with a size in $R-\phi \times z$ of $50 \times 400\ \mu\text{m}^2$ [11]. It consists of three concentric cylinders in the barrel part and three end-caps disks at each side. The SCT provides additional information about the trajectory of the particles and also contributes to determining the momentum and position of the vertex. The detection area consists of four cylindrical barrel layers and nine end-caps on each side. Each of the 4088 silicone detectors located on all SCT layers contains 780 readout strips on both sides. More than 6 million readout channels ensure high accuracy in determining the position of charged particles.

The TRT, a straw-tube tracker, is the upper part of the ID. It provides a large number of tracking points and consists of drift tubes with a diameter of 4 mm [12]. Each tube has in its center gold-plated tungsten wire and acts as a small proportional counter. The barrel region of TRT contains 52 544, 1.5 m long,

straw tubes parallel to the beam axis. Wires in these straws are divided into two halves and read out at both ends of the straw. Each side of end-caps consists of 122 880 radial 0.4 m long straws. Thanks to the emission of the transition radiation, the TRT is used to determine the type of particles that have flown through the detector, such as electron or pion (electron much more likely emits transition radiation than pion of the same momentum).

2.3.3 Calorimetry

Another part of the ATLAS detector is the calorimeter system. Calorimeters are used to determine the energy of particle shower. There are two basic types of particle showers: electromagnetic and hadronic. Electromagnetic showers are mostly electrons and photons interacting with matter via electromagnetic force. Hadrons produce hadronic showers which proceed mostly via strong nuclear force. In most cases, the calorimeter is designed in the way that the particles are absorbed and transfer all their energy within the detector. Calorimeters are able to stop most particles, except muons and neutrinos.

Figure 2.4 shows the calorimeter system of the ATLAS detector. This system consists of two calorimeters, namely Liquid Argon Calorimeter (LAr) and Tile Calorimeter (TileCal). Together they cover the pseudorapidity range $|\eta| < 4.9$ and are symmetrical along the azimuth angle.

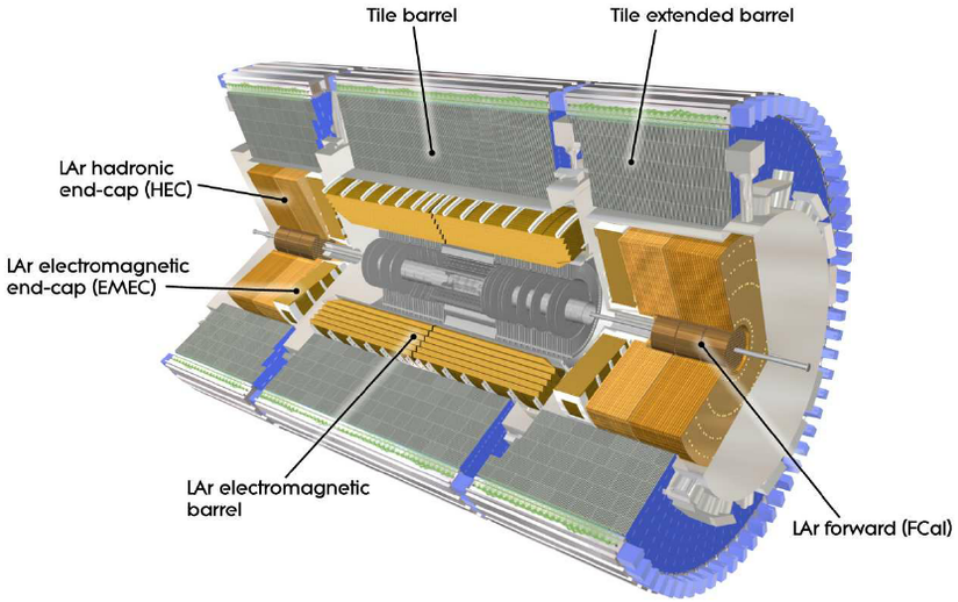


Figure 2.4: Cut-away view of the ATLAS calorimeter system [10]

The LAr is located around the inner detector. This calorimeter consists of an electromagnetic barrel ($|\eta| < 1.475$) and end-caps, which are divided into forward calorimeter (FCal) ($3.1 < |\eta| < 4.9$), electromagnetic (EMEC) ($1.375 < |\eta| < 3.2$) and hadron (HEC) end-caps ($1.5 < |\eta| < 3.2$). It uses liquid argon and lead (barrel, EMAC) or copper (HEC, FCal) as an absorber as active material. TileCal

($|\eta| < 1.7$) is located behind the electromagnetic calorimeter and uses scintillating tiles as the active material and steel as the absorber. The electromagnetic calorimeter is ideal for precision measurements of particles that interact electromagnetically, such as electrons or photons, thanks to its fine granularity. The rest of the calorimeter has a coarser granularity; nevertheless, it satisfies the physical requirements for jet reconstruction and missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ measurements.

Electromagnetic Calorimeter

Each of the EM calorimeter components is housed in their separate cryostat. The barrel is placed in one cryostat with the central solenoid and EMECs share cryostat with FCal and HECs. EM calorimeter's kapton readout electrodes are placed in gaps between the accordion-shaped lead absorber plates. This accordion geometry provides complete azimuthal symmetry. The barrel is divided into two identical halves at $z = 0$, and each end-cap consists of two coaxial wheels.

Hadronic End-cap Calorimeter

The HEC is a sampling calorimeter and consists of two separate wheels per end-cap located directly behind EMEC. Each wheel is made out of 32 identical wedge-shaped modules and divided into two layers in depth. Readout kapton electrodes are located in gaps between copper plates that are perpendicular to the beam axis.

Forward Calorimeter

The FCal is placed between the HEP and the beam pipe and consists of three modules for each end-cap. The first module, FCal1, is made of copper and is designed for electromagnetic measurements. The other two models, FCal2 and FCal3, made of tungsten, are optimized for the measurement of energy in hadronic interactions. Each model consists of regularly spaced longitudinal channels in the metal matrix. These channels are filled with the electrode structure built from concentric rods and tubes parallel to the beam axis.

Tile Calorimeter

The TileCal is placed around the electromagnetic calorimeter. Together with an electromagnetic calorimeter, TileCal provides highly accurate measurements of hadrons, jets with $|\eta| < 1.6$ and $E_{\mathrm{T}}^{\mathrm{miss}}$. It consists of barrel, that covers region $|\eta| < 1.0$ and extended barrel on each side, covering the region $0.8 < |\eta| < 1.7$. Each component is divided azimuthally into 64 wedge-shaped modules and three layers in depth. The layers are further split into cells.

TileCal uses iron plates as absorbers and plastic scintillation tiles as the active material. The scintillation tiles are placed in a plane perpendicular to the particle

beams. The output from the scintillation tiles (scintillation radiation) passes through wavelength-shifters into a photomultiplier tube, which converts the light signal into an electric pulse.

2.3.4 Muon System

In calorimeters, all known particles except muons and neutrinos are absorbed. The muon spectrometer (MS) is used to detect muons that penetrate through all the detectors or, more precisely, to determine their momentum and track. MS is the outermost part of the ATLAS detector and therefore covers the largest area. It has four subsections: Monitored Drift Tubes (MDTs), Cathode Strip Chambers (CSCs), Resistive Plate Chambers (RPCs), and Thin Gap Chambers (TGCs), see Fig. 2.5. All of them are placed inside the magnetic field generated by the large barrel toroid and two end-cap magnets placed by both ends of the barrel toroid. Muon tracks in the region $|\eta| < 1.4$ are bend by toroid, which generates the magnetic field of $B = 1.5$ T. For $1.6 < |\eta| < 2.7$ magnetic deflection is caused by end-caps, that generate the magnetic field of $B = 1$ T. In transition region, $1.4 < |\eta| < 1.6$ magnetic bending is provided by a combination of both fields.

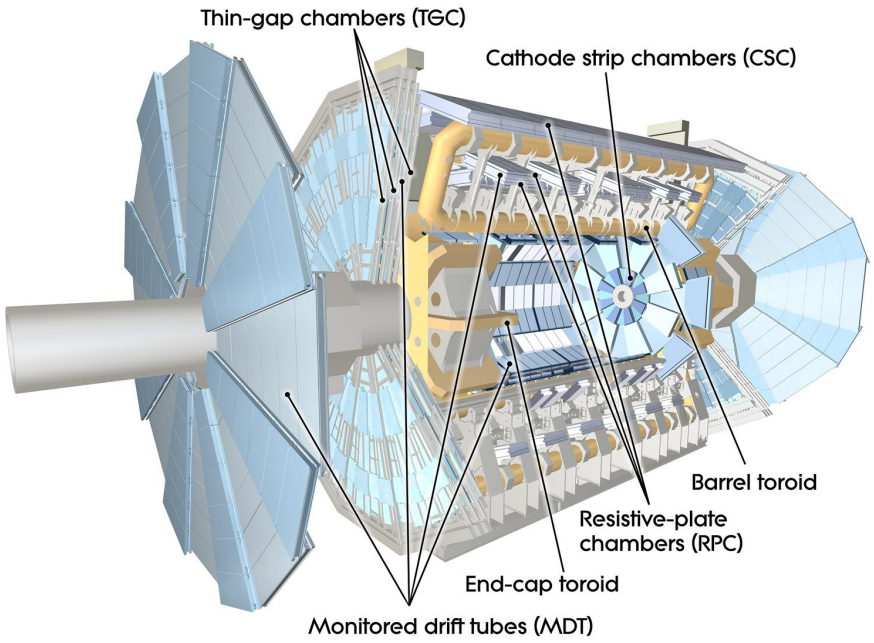


Figure 2.5: Cut-away view of the ATLAS muon spectrometer [10]

Muon tracks are measured in 4000 individual muon chambers arranged in three cylindrical layers in the barrel region and three layers of planes perpendicular to the beam axis in the transition and end-cap region. A precision measurement of the track coordinates in the bending direction is provided by MDTs ($|\eta| < 2$) and CSCs ($2 < |\eta| < 2.7$).

The trigger system for the MS covers the region $|\eta| < 2.4$ and consists of RPCs in the barrel and TGCs in the end-cap regions. All the chambers provide well-

defined p_{T} thresholds and bunch-crossing identification and also measure the muon track coordinates in the direction orthogonal to the bending one.

2.3.5 Trigger and Data Acquisition System

The LHC is designed to collide proton bunches every 25 ns, corresponding to the bunch-crossing rate of 40 MHz [13]. In Run 2, instantaneous luminosity peaked at 21 $\times 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and this lead to a high pile-up of multiple pp interactions per bunch-crossing ($\langle\mu\rangle = 33.7$ for Run2 [14]). All this resulted in large and complex data flows with a data volume of more than 60 terabytes per second. For discovering new physics, it would be beneficial to store all the data, but storing event data is limited to at most a few GB per second. Therefore, a huge rejection factor is desired. Luckily, not all events contain interesting characteristics for physics discoveries. These interesting characteristics are, e.g., high-p_{T} electrons, photons, jets, τ-leptons, large $E_{\mathrm{T}}^{\mathrm{miss}}$, or total transverse energy. Here comes the trigger and data acquisition system. It is responsible for selecting and storing events of interest for further study. Events from ATLAS are selected with a two-level trigger system, hardware-based first level (L1), and a software-based high-level trigger (HLT).

The L1 uses custom-made electronics to determine Regions-of-Interest (RoIs) (a region in $\eta \times \mu$ where interesting features are found) retrieved from coarse granularity calorimeter (L1Calo) and muon detector (L1Muon) and to be investigated by HLT. L1 reduces the event rate from LHC bunch-crossing rate of 40 MHz to 100 kHz, and the decision to keep the data from the event is made within 2.5 μs after the event occurs [15]. The decision from L1 is formed by the Central Trigger Processor (CTP). In upgrade before LHC Run 2 new component, the topological trigger (L1Topo) was included, where topological requirements are applied to geometric or kinematic combinations between trigger objects received from L1Calo or L1Muon.

The HLT runs offline-like algorithms executed on a large computing farm of approximately 40 000 processor cores, designed to make decisions typically within 300 ms. The HLT usually provides 2500 independent trigger chains, and these sequences of algorithms execute a feature-extraction algorithm, which requests event-data fragments from within the L1 RoIs, and the hypothesis algorithm decides whether or not to keep data. Events that pass this sequence are written into a different data stream. To avoid unnecessary consumption of the available bandwidth, either full event or only partial event information is written out, depending on further usage, e.g., for physics analysis (full event), detector calibration, monitoring, or trigger level analysis (partial event information).

2.4 Object Reconstruction on the ATLAS

2.4.1 Jet Reconstruction and Calibration

The reconstruction of particle showers, implemented in the topological clustering algorithm, is possible thanks to the lateral and longitudinal segmentation of the calorimeters. The three-dimensional topo-clusters [16] are formed by topologically connected calorimeter cells whose absolute energy measurements significantly exceed the expected noise. The energy of a topo-cluster is given by the sum of energy of all included cells, and its mass is set to zero. The position is determined from the energy-weighted mean of the cell position in $\eta \times \phi$. For jet reconstruction, only topo-clusters with positive energy are used.

LCTopo Jets

Topo-clusters can be calibrated in two ways: calibration to the electromagnetic scale (EM) and with local calibration (LC) weighting [17]. The EM topo-clusters are calibrated to the response from electrons, while the LC topo-clusters are classified as electromagnetic or hadronic. Then a correction by weighting scheme is applied for the different electron-to-pion response in the calorimeters, and correction for non-compensation, dead material, and out-of-cluster energy deposits is used. LCTopo jet reconstruction uses the anti-k_{t} algorithm with radius parameter of $R = 0.4$ and input data are topo-clusters at LC scale.

Particle flow jets

The topo-clusters and tracks form the basic inputs for the particle flow algorithm [18] [19]. The particle flow algorithm determines a list of the topo-clusters containing unchanged topo-clusters and also a set of new topo-clusters, resulting from the energy subtraction procedure and a list of tracks. The first step in this algorithm is applying 'tight selection' and additional criteria on well-measured tracks. Selected tracks are then attempted to be matched to a single topo-cluster in the calorimeter. Based on the track momentum and the topo-cluster position, the algorithm computes the expected energy deposited in the calorimeter for the particle that also created the track. The next step evaluates the probability that the particle deposited energy in more than one topo-cluster (it is relatively common). The algorithm then adds topo-clusters to the track/topo-cluster system if it is necessary for recovering the full shower energy. The expected energy is subtracted cell by cell from the set of matched topo-clusters, and the remnants are removed if the remaining energy can be matched with the expecting shower fluctuations of a signal for a single particle.

Particle flow (PFlow) jet reconstruction uses the anti-k_{T} algorithm with the radius parameter of $R = 0.4$. The input data to jet reconstruction are the set of topo-clusters at EM scale that survived the energy subtraction step and the selected tracks matched to the hard-scatter primary vertex.

Jet Calibration

The jet calibration procedure consists of origin correction, pile-up correction, numerical inversion based on MC simulation correcting detector response, global sequential correction process for correction of additional fluctuations in jet response and residual in situ calibration. For PFlow jets, no origin correction is needed as topo-cluster η and ϕ are recomputed with respect to the primary vertex position before jet reconstruction.

The jet origin correction forces the jet to point back to the primary vertex rather than to the center of the detector while keeping the jet energy constant. During Run 2 at center-of-mass energy in pp collisions $\sqrt{s} = 13$ TeV and with a reduced bunch spacing interval, the pile-up increased and therefore in the next step of the jet calibration, the energy transferred from the pile-up interactions is removed from the jet. Subsequently, the jet energy scale and η calibration factors are applied, which adjust the jet energy in the EM/LC scale to the particle-level energy scale. These factors are obtained using events from the MC simulation. Therefore, since the jet calibration is dependent on the MC, additional correction of the differences between the data and the MC is required. This is removed by a residual in situ calibration applied to the data only. The transverse momentum balance of a jet and a well-measured reference objects is used for this correction. $Z(ee)$ + jet and γ + jet events are used to calibrate jets with p_{T} up to 1 TeV. For jets with higher p_{T} , multijets are used, i.e. events with large p_{T} jet and jets with much lower p_{T} flying to the opposite side in transverse plane, and therefore it is possible to calibrate all these jets using Z/γ + jet events. In addition to all this, dijet events (dijet balance technique) is used to uniformize the jet response in all $|\eta|$ areas. Jets from a lower $|\eta|$-area are used to calibrate jets with a higher $|\eta|$.

2.4.2 Tau Reconstruction and Identification

As mentioned in Chapter 1, the tau leptons can decay either leptonically ($\tau \to \nu_\tau \nu_l l, l = e, \mu$ or hadronically ($\tau \to \nu_\tau$ hadrons) and due to its short decay length, they typically decay before reaching active regions of the ATLAS detector, and therefore, can only be identified via their decay products. In this chapter, we will only take into account hadronically decaying taus $\tau_{\mathrm{had\_vis}}$ (neutrinos can not be reconstructed). Since 68% of hadronic tau leptons decays include at least one neutral pion, and almost all of them include one (72%) or three charged hadrons (22%), their signature in the detector corresponds to the signature of a narrow jet with one or three tracks.

The $\tau_{\mathrm{had\_vis}}$ reconstruction algorithm is seeded by LCTopo jets described above. Input data for this algorithm are LC topo-clusters. Jet must additionally fulfill the requirement to have $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.5$ to seed a $\tau_{\mathrm{had\_vis}}$ candidate.

Every event must have a reconstructed primary vertex with at least three associated tracks. There can be multiple primary vertex candidates, and from these candidates, the tau lepton production vertex is identified. The largest frac-

tion of the p_{T} sum of all tracks with $p_{\mathrm{T}} > 0.5$ GeV within an angular distance $R = 0.2$ around the seed jet axis (energy-weighted barycentre of all associated topo-clusters) is matched to the primary vertex candidate, which is subsequently chosen as the tau vertex. With the use of the tracks' p_{T}, the number of hits in the tracking detectors, and other impact parameters with respect to the tau vertex, a set of boosted decision trees (BDTs) classify all tracks within $R = 0.4$ around the tau direction as *core* ($R < 0.2$) and *isolation* ($0.2 < R < 0.4$) tracks. The number of core tracks defines the number of charged particles (prongs). Secondary vertex is constructed for all core tracks, in case that $\tau_{\mathrm{had\_vis}}$ candidate have more than one core track.

The energy of the $\tau_{\mathrm{had\_vis}}$ is obtained through dedicated calibration schemes described in detail in Ref. [20]. For a baseline correction, first, a correction to the sum of the energy of the LC topo-clusters within $R = 0.2$ of the tau candidate is applied for energy contribution from the pile-up. The information from the baseline combined with the "Tau Particle Flow" method described in Ref. [21] in a boosted regression tree is used to calculate the final energy of the $\tau_{\mathrm{had\_vis}}$ candidate.

The tau reconstruction algorithm does not reject other particles that result in a jet-like signature. To identify hadronic tau lepton decays in data recorded in Run2, tau identification algorithm based on Recurrent Neural Network (RNN) [22] was used. It uses track and calorimeter information in an RNN to reject background from the quark- and gluon-initiated jets. The RNN for tau candidates associated with one and three tracks are trained separately using a simulated signal sample consisting of $\gamma^* \to \tau\tau$ events and dijet events for background. Observables used as input in the different parts of the RNN are the following: cluster depth ($\lambda_{\mathrm{cluster}}$), longitudinal cluster extension ($\langle\lambda^2_{\mathrm{cluster}}\rangle$), $p_{\mathrm{T}}^{\mathrm{seedjet}}$, $p_{\mathrm{T}}^{\mathrm{uncalib}}$, central energy fraction (f_{cent}), mass of the track system (m^{track}), mass of the track-plus-EM-system ($m^{\mathrm{EM+track}}$), and many others. Four working points (WPs) (*Very loose, Loose, Medium, and Tight*) corresponding to signl efficiencies are provided and summarized in Tab. 2.1.

Table 2.1: List of defined working points. Corresponding RNN cuts are: RNN score $> 1-$ signal efficiency

	Signal efficiency	
Working point	1-prong	3-prong
Tight	60%	45%
Medium	75%	60%
Loose	85%	75%
Very Loose	95%	95%

2.4.3 Muon Reconstruction

At first, muon reconstruction [23] is carried out independently in ID and MS. Subsequently, the muon track is formed from the combined information obtained from ID and MS. The muons are divided into four groups, depending on which of the subdetectors was used for reconstruction: segment-tagged muons, calorimeter-tagged muons, extrapolated muons, and combined muons (CDs), which cover the most significant fraction and are used in the following analysis.

For the ID muons, the reconstruction proceeds in the same way like for any other charged particle. MS muon reconstruction, first, forms segments by a search for hit patterns inside each muon chamber. The segment is a straight line fitted to the hits found in each layer of the MS, and the hits must be aligned on a trajectory in the bending plane of the detector. Muon track candidates are formed by fitting together hits from segments in different layers. The CB muons are reconstructed following an outside-in pattern, which means that matching starts with the muons tracks to be reconstructed in the MS and then extrapolated inward to the center of the detector and matched to an ID track. The reconstructed muon track is obtained by fitting hits formed simultaneously in both the ID and MS.

For muon identification, the quality requirements that suppress background, are applied, while selecting muons from W, Z, H decays or from decays of new unknown particle (prompt muon), instead of muons from decays of hadrons with high efficiency and/or guaranteeing a robust momentum measurement. Four identification working points marking muon quality are provided *Loose, Medium, Tight* and *High-*p_{T} .

2.4.4 Electron Reconstruction and Identification

Electrons are reconstructed from the information measured by the ID and the EM calorimeter [24]. The first step in electron reconstruction is localizing cluster seeds of energy deposits found within the EM calorimeter. The seeds are identified from localized energy deposit using a sliding-window algorithm of size 3×5 towers in $\eta \times \phi$ with steps of 0.0025 in either the η or ϕ direction and searches for local maxima with $E_{\mathrm{T}} > 2.5$ GeV. The next step is identifying charged-particle tracks in the ID. Sets of three space-points in the silicon-detector layers form the track seeds and track reconstruction then proceeds in pattern recognition and track fitting. In the pattern recognition, a pion hypothesis for the model of energy loss is used in a first attempt to extend a track seed with $p_{\mathrm{T}} > 1GeV$ into a full track of at least seven hits in the silicon detectors. If the extension algorithm fails, a second attempt is made with the electron hypothesis if the EM cluster satisfies requirements on the shower width and depth. Then, the tracks are fitted according to the hypothesis used in the pattern recognition. If a track fit fails under the pion hypothesis it is refit with the electron hypothesis. The third step is matching EM clusters with tracks. This matching takes into account η and ϕ coordinates, energy-loss due to bremsstrahlung and the number of precision hits in the silicon detector. Finally, the cluster size is enlarged to 3×7 (5×5) in the barrel (end-cap) EM calorimeter region and used for reconstruction of the

clusters formed around the seed clusters.

For Run 2 data analyses, the likelihood-based (LH) method was used as he baseline electron identification algorithm [25]. The LH method is multivariate analysis (MVA) technique which evaluates multiple properties of the electron candidate while making a selection decision. Identification operating points provided for electron identification, in order of increasing background rejection, are LH*Loose*, LH*Medium*, LH*Tight*.

3. Data and Monte Carlo samples

3.1 The Data

For the purposes of this master thesis, the data collected in Run 2 (2015-2018) were used. During this period, the center-of-mass energy of the proton-proton collisions reached the value $\sqrt{s} = 13$ TeV. The data correspond to the total integrated luminosity 138.8 fb^{-1}.

The quality of the data is identified as good, as during the period of data collection for our analysis events were collected with stable beams only. The events passed the data quality requirements of the ATLAS subdetectors used for the reconstruction. In addition, there has to be more then one track associated to the hardest primary vertex in the event identified as the one with highest $\sum p_{\mathrm{T}}^2$ of associated tracks.

3.2 The Monte Carlo samples

Monte Carlo samples for $Z(\mu\mu)$ + jets were simulated at next-to-leading order (NLO) using Powheg-Box [26, 27, 28] interfaced to Pythia 8 [27] and the CT10 [29] parton distribution functions (PDF). Simulation of final-state photon radiation in Z boson decays was performed by the Photos [30] package. A requirement on the minimum $p_{\mathrm{T}} > 4$ GeV of each muon is applied at the generator level.

The $t\bar{t}$ and single-top event samples were generated with Powheg-Box v2 [31, 32, 33] interfaced with Pythia 8.230 [34] for the parton shower and hadronization, using the A14 tune [35] and the NNPDF2.3lo PDF set [36]. The Powheg-Box v2 provides matrix elements at NLO in α_{S} with the NNPDF3.0nlo PDF set [37]. The NLO radiation factor h_{damp} was set to 1.5 times the mass of the top quark m_{top}. For the removal of the interference between $t\bar{t}$ production and Wt-channel single-top production, the diagram removal method was used [38].

The diboson samples were simulated at next-to-next-to-leading order (NNLO) using Sherpa 2.2.1 [39] with the NNPDF3.0nnlo PDF set. The matrix elements were calculated for up to one parton at NLO and three partons at LO using Comix [40] and OpenLoops [41, 42, 43] matrix-element generators, and merged with the Sherpa parton shower [44] using ME+PS@NLO [45] prescription.

All simulated event samples were passed through the simulation of the ATLAS detector based on Geant4 [46, 47] and were reconstructed using the same softwares as used for the data. The effect of the multiple pp interactions in the same and neighboring bunch crossing was included by overlaying simulated minimum-bias events onto the original event. The MC samples were reweighed so that the distribution of the average number of pile-up interactions matches to the data.

4. Fake Factor Method

Misidentification significantly contributes to the background for the physics analyses using hadronic τs. Thanks to the τ identification of the ATLAS detector, the physics signatures with τ in the final state have large background rejection. The majority of jets is suppressed by τ identification criteria. Despite this fact, significant number of misidentified (fake) $\tau_{\text{had\_vis}}$ objects remain and that is due to large cross section of the QCD jet production at the LHC relative to that of τ production. Monte Carlo simulation is not able to well model the fake τ background, thereupon data-driven methods had to be developed. In the following sections, the data-driven fake factor (FF) method [48][49] will be discussed.

4.1 Introduction to the Fake Factor Method

The fake factor method's basic concept is straightforward: selecting a signal region with requirement on failing the RNN identification criteria (fail-ID SR) that is enriched in fake τ's, and then using an extrapolation factor to estimate the fake τ background in the certain signal region (pass-ID SR). This method is data-driven, which means that the fail-ID SR is selected in data. The extrapolation factors are referred to as the fake factors (FFs) and are measured in dedicated CRs selected in data. The CRs are chosen so that as many selection criteria as possible were the same as for the SR, and several of them differed so that the CR was enriched in fake-τs. In most cases, a minimum cut is imposed on the RNN score in fail-ID region. The FF is defined as the ratio of the number of events with jets reconstructed as $\tau_{\text{had\_vis}}$ candidates - fake-τs that pass all selecting criteria in the CR ($N^{\text{CR}}_{\text{passID}}$) to the number of the events failing the RNN identification criteria for given working point ($N^{\text{CR}}_{\text{failID}}$):

$$\text{FF} = \frac{N^{\text{CR}}_{\text{passID}}}{N^{\text{CR}}_{\text{failID}}}. \tag{4.1}$$

The FFs are usually measured in bins of p_{T} , η and number of associated tracks in the τ hadronic decay (1-prong, 3-prong) or different decay modes. Assuming the defined SR requires $\tau_{\text{had\_vis}}$ object to pass the identification criteria of RNN working point, the number of events that contain fake-τ is estimated as:

$$N^{\text{SR}}_{\text{passID}} = N^{\text{SR}}_{\text{failID}} \times \text{FF}, \tag{4.2}$$

where $N^{\text{SR}}_{\text{failID}}$ corresponds to the number of events that contain $\tau_{\text{had\_vis}}$ candidate in SR-like region failing the RNN working point. The SR-like region is defined in the same way as the nominal SR with the exception of failed identification criterion on RNN working points. The CR and fail-ID SR can contain some true hadronic taus so in order to obtain $N^{\text{SR}}_{\text{failID}}$, $N^{\text{CR}}_{\text{failID}}$ and $N^{\text{CR}}_{\text{passID}}$, contribution from events with true hadronic taus in either region need to be subtracted using the MC simulation. The Eq. 4.1 and Eq. 4.2 can be rewritten as:

$$\mathrm{FF} = \frac{N^{\mathrm{CR}}_{\mathrm{passID}} - N^{\mathrm{CR,MC}}_{\mathrm{passID,true}\tau}}{N^{\mathrm{CR}}_{\mathrm{failID}} - N^{\mathrm{CR,MC}}_{\mathrm{failID,true}\tau}}, \tag{4.3}$$

$$N^{\mathrm{SR}}_{\mathrm{passID}} = (N^{\mathrm{SR}}_{\mathrm{failID}} - N^{\mathrm{SR,MC}}_{\mathrm{failID,true}\tau}) \times \mathrm{FF}, \tag{4.4}$$

The jets initiated by quarks and gluons differ in their width (see. Fig 4.1), gluon-initiated jets tend to contain more particles and be wider than those initiated by quarks. This suggests that the probability to be reconstructed as a $\tau_{\mathrm{had\_vis}}$ and therefore also FFs will differ for quark- and gluon-initiated jets. Hence, the CR should have similar quark/gluon fraction to that in SR.

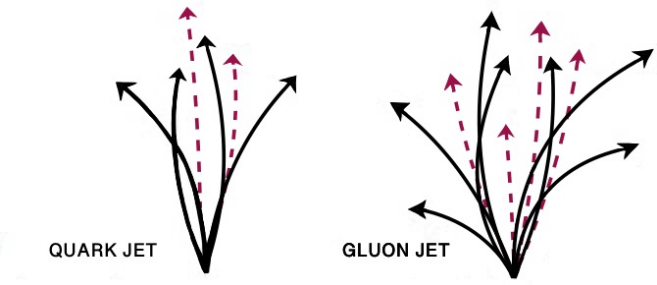


Figure 4.1: Illustration of the difference between quark- and gluon-initiated jets. The jets initiated by quarks and gluons differ in their width, gluon-initiated jets tend to contain more particles and be wider than those initiated by quarks.

4.2 The Universal Fake Factor Method

In practice, it is difficult to ensure that the CR will have similar quark/gluon composition to that of SR. Nevertheless, we are developing a new universal FF method for estimating fate-τ background on ATLAS. This universal FF method accounts for the quark/gluon composition in any particular SR. Taking into account only quark and gluon-initiated jets, the number of fake-τs can be calculated by combined FF for quark- and gluon-initiated jets. The Eq. 4.2 for number of events containinh the fake-τs can be written as:

$$\begin{aligned} N^{\mathrm{SR}}_{\mathrm{passID}} &= \mathrm{FF}_q \times N_q + \mathrm{FF}_g \times N_g \\ &= N^{\mathrm{SR}}_{\mathrm{failID}} \times \mathrm{FF}_{comb}, \end{aligned} \tag{4.5}$$

where FF_{comb} can be defined as function of the quark fraction:

$$\mathrm{FF}_{comb} = q\mathrm{FF}_q + (1-q)\mathrm{FF}_g, \tag{4.6}$$

with q denoting the fraction of quark-initiated jets in the fail-ID SR and FF_q, FF_g corresponds to the FFs measured in pure quark and pure gluon sample, respectively. In Fig. 4.2 the FF is plotted as a function of q.

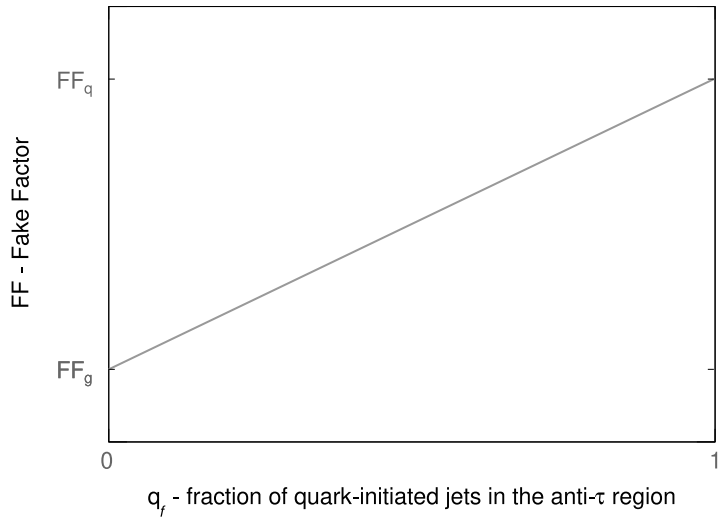


Figure 4.2: The linear dependence of FF on the quark fraction q

It is unlikely to find pure quark or gluon sample, and eventually measure FF_q and FF_g. However the connection between FF and q can be determined by interpolation. The FFs measured in two discretionary regions can be defined as:

$$\mathrm{FF}_1 = q_1\mathrm{FF}_q + (1-q_1)\mathrm{FF}_g, \tag{4.7}$$

$$\mathrm{FF}_2 = q_2\mathrm{FF}_q + (1-q_2)\mathrm{FF}_g, \tag{4.8}$$

and can be measured in the usual way given by the Eq. 4.1:

$$\mathrm{FF}_1 = \frac{N^{\mathrm{CR}_1}_{\mathrm{passID}} - N^{\mathrm{CR}_1,\mathrm{MC}}_{\mathrm{passID,true}\tau}}{N^{\mathrm{CR}_1}_{\mathrm{failID}} - N^{\mathrm{CR}_1,\mathrm{MC}}_{\mathrm{failID,true}\tau}}, \tag{4.9}$$

$$\mathrm{FF}_2 = \frac{N^{\mathrm{CR}_2}_{\mathrm{passID}} - N^{\mathrm{CR}_2,\mathrm{MC}}_{\mathrm{passID,true}\tau}}{N^{\mathrm{CR}_2}_{\mathrm{failID}} - N^{\mathrm{CR}_2,\mathrm{MC}}_{\mathrm{failID,true}\tau}}, \tag{4.10}$$

Using the above system of equations 4.7, 4.8, FF_q and FF_g can be written as:

$$\mathrm{FF}_q = \frac{(1+q_2)\mathrm{FF}_1 - (1+q_1)\mathrm{FF}_2}{q_2 - q_1}, \tag{4.11}$$

$$\mathrm{FF}_g = \frac{q_2\mathrm{FF}_1 - q_1\mathrm{FF}_2}{q_2 - q_1}, \tag{4.12}$$

therefore for the measuring of FF_q and FF_g is necessary to know the quark fractions q_1 and q_2.

The procedure of acquiring the dependance of FF on quark fraction in anti-ID region by analytically computing FF_q and FF_g is shown in Fig. 4.3.

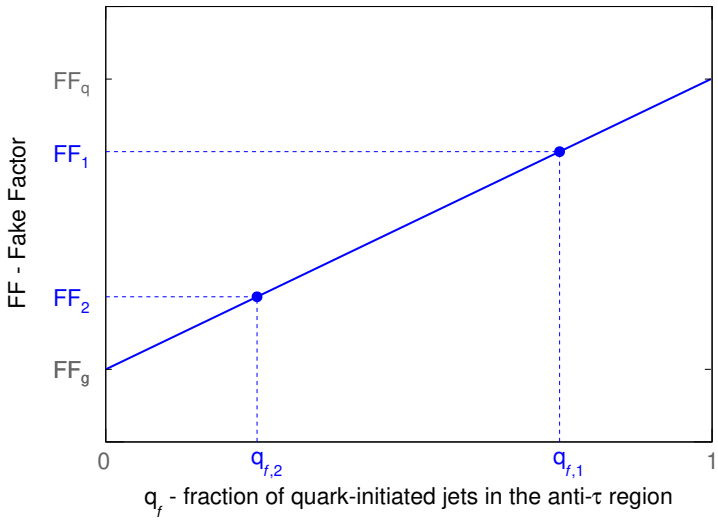


Figure 4.3: The linear dependence of FF on the quark fraction q determined by analytically computing FF_q and FF_g

Now, having the fail-ID SR, we are getting back to Eq. 4.6 and the FF can be interpolated as:

$$\mathrm{FF}_{SR} = q_{SR}\mathrm{FF}_q + (1 - q_{SR})\mathrm{FF}_g, \tag{4.13}$$

where FF_q is given by Eq. 4.11 and FF_g by Eq. 4.12. The interpolation of FF_{SR} using pure quark FF_q and pure gluon FF_g is illustrated in Fig. 4.4.

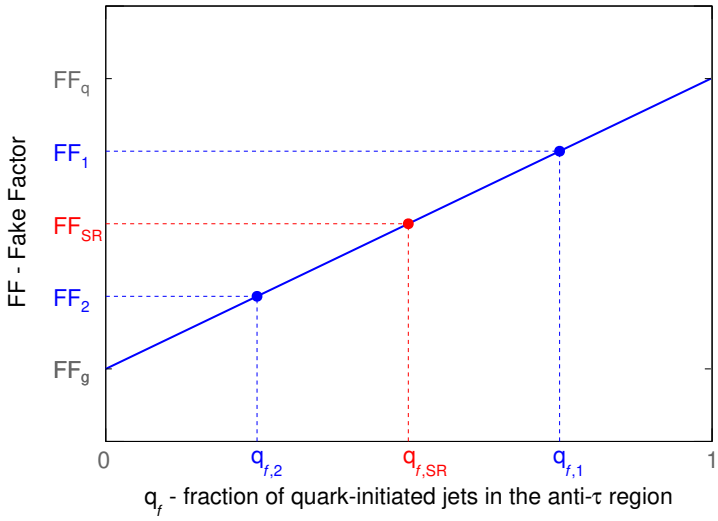


Figure 4.4: The interpolation of FF_{SR} using pure quark FF_q and pure gluon FF_g

The quark fractions of interest q_1, q_2 and q_{SR} cannot be estimated form data directly, therefore a template fit method is used. The quark and the gluon template

can be obtained by truth-matching quark- and gluon-initiated jets in MC as a function of a variable that can distinguish well between quarks and gluons. Such variable is the jet width, which is p_{T} -weighted ΔR (the distance of constituent's direction from the jet axis) of constituents associated to the jet:

$$j = \frac{\sum_i \Delta R^i p_{\mathrm{T}}^i}{\sum_i p_{\mathrm{T}}^i}, \tag{4.14}$$

where the index, i, refers to the i^{th} constituent in the reconstructed jet.

The track-based jet width is calculated using JetTrackMomentsTool [50]. The calculation of track-based jet width is loop over tracks with $p_{\mathrm{T}} > 1$ GeV associated to the primary vertex of seed jet. If no tracks fulfill the requirements, the track-based jet width is set to "-1".

5. The FF measurements

In the previous chapter, we described the universal fake factor method. In this method, from measured FF_1 and FF_2, the FF in any particular SR can be interpolated. These FFs should be measured in the regions that are dominant in the production of either quark- or gluon-initiated jets. For gluon-initiated jets, the corresponding region is the multi-jet region and for quark-initiated jets, it is the Z+jets region, and this master thesis focuses on the FF measurement specifically in this region. The Z boson associated with jet is produced in process $q+g \to Z+q$ with corresponding leading order Feynman diagram shown in Fig.1.2a.

5.1 Events and Object Selection

The Z+jets region consists of $Z(\mu\mu)$ events. The following object and events selection is summarized in Tab.5.1. The data are saved in Analysis Object Data (xAOD) format. This format contains all events, which causes that it is too big and, moreover, lacks detailed information on physics objects needed by Combined Performance (CP) tools. Thus, there is a need for a derived format (DxAOD). The derivations for our analysis were performed by TAUP3. The TAUP3 is one of the tens of derived data formats. It contains events with at least one muon and one tau candidates, and therefore there is a big reduction in the number of events.

Some of the muons in events can originate from b- and c-quarks decaying inside jets. For suppression of muons from jets, muon isolation is applied by the trigger. The muon isolation requires a low p_{T} (sum of p_{T} carried by particles in the cone) in a cone around the muon direction. The single-muon trigger chain HLT_26_ivarmedium was used. It selects isolated muons with the p_{T} threshold of 26 GeV[51]. It is ideal to use a trigger chain with the lowest p_{T} threshold, but a trigger with such a low p_{T} threshold as 26 GeVwould quite possibly overload the trigger system, and therefore the cut on muon isolation is required. However, the muon isolation has finite efficiency, so where possible (i.e. for $p_{\mathrm{T}} > 50$ GeV), we help with an uninsulated trigger HLT_mu50, designed to collect muons with large p_{T} without efficiency loss due to any isolation requirement. For 2015, the trigger chain HLT_mu20_iloose_L1MU15 (p_{T} threshold of 15 GeVfor L1 trigger and 20 GeVfor HLT) was used instead of HLT_26_ivarmedium.

Jets are reconstructed using the anti-k_{t} algorithm with a radius parameter of $R = 0.4$ and with particle-flow objects as inputs, and have $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 4.5$. After this selection, the OverLapRemoval (OLR) is applied. It makes sure that each object is classified as just one object type. A jet is ignored within $\Delta R = 0.2$ of an electron or hadronically decaying τ candidate.

The $\tau_{\mathrm{had\_vis}}$ candidates must have $p_{\mathrm{T}} > 20$ GeV, $|\eta| > 1.37$ or $1.52 < |\eta| < 2.5$, one or three associated charged tracks, and an absolute electric charge of one. No RNN cuts applied. In OLR, the $\tau_{\mathrm{had\_vis}}$ candidates are ignored if they are within $\Delta R = 0.2$ of a muon or electron candidate.

The electron candidates are required to pass LHLoose identification selection, to have $p_{\mathrm{T}} > 15$ GeV and $|\eta| < 2.47$. The electron candidates have to be compatible with originating from the primary vertex by satisfying $|d_0|/\sigma_{d_0} < 5$ and $|z_0 \sin\theta| < 0.5$ mm, where $d_0(z_0)$ is the transverse(longitudinal) impact parameter relative to the primary vertex, and σ_{d_0} is the uncertainty in d_0. Electron candidates are ignored if they share their reconstructed track with a muon candidate or if their angular distance from a jet is within $0.2 < \Delta R < 0.4$. The electron candidates should be good electrons, meaning they are not the noise from the EM calorimeter.

Muon candidates are required to pass Loose Quality identification criteria and to have $p_{\mathrm{T}} > 7$ GeV and $|\eta| < 2.5$. To ensure that muon candidates originate from the primary vertex, they must possess $|d_0|/\sigma_{d_0} < 5$ and $|z_0 \sin\theta| < 0.5$ mm. Muon candidates are ignored if their angular distance from a jet is $\Delta R < 0.4$ with the following exceptions: If $\Delta R < 0.2$ or the muon track is associated with the jet, and if the jet has either less than three tracks or less than twice the transverse momentum of the muon candidate, the jet is removed instead.

Table 5.1: Summary of object and events selection

Trigger	
2015	HLT_mu20_iloose_L1MU15 OR HLT_50
2016-2018	HLT_mu26_ivarmedium OR HLT_50
Object selection before OLR	
jets	PFlow, $p_{\mathrm{T}} > 20$GeV, $\|\eta\| < 4.5$
taus	$p_{\mathrm{T}} > 20$GeV, 1 or 3 tracks, $\|\eta\| < 1.37$ OR $1.52 < \|\eta\| < 2.5$, no RNN cut
electrons	ID LHLoose, $p_{\mathrm{T}} > 15$GeV, $\|\eta\| < 2.47$, $\|d_0\|/\sigma_{d_0} < 5$, $\|z_0 \sin\theta\| < 0.5$, good electrons
muons	Quality Loose, $p_{\mathrm{T}} > 7$GeV, $\|\eta\| < 2.5$, $\|d_0\|/\sigma_{d_0} < 5$, $\|z_0 \sin\theta\| < 0.5$
Event selection after OLR	
$m_{\mu\mu} \in (70, 110)$GeV	
Exactly two muons and zero electrons surviving the OLR	
At least one tau surviving the OLR	
Muon quality cut on the two muons: Medium	
Muon isolation cut on the two muons: FCTightTrackOnly	
Leading muon $p_{\mathrm{T}} > 27.3$GeV	
Subleading muon $p_{\mathrm{T}} > 10$GeV	

In $Z(\mu\mu)$+jets events selection, at least one tau (fake-tau candidate), precisely two muons, and zero electrons surviving the OLR are required since there are exactly two muons in the final state of Z decay and no electron. The invariant mass of these two muons $m_{\mu\mu}$ is required to be between 70 GeV to 110 GeV

to ensure they come from the decay of the Z boson. Both muons are required to pass Medium Quality identification criteria. In addition, muons must satisfy the 'FCTightTrackOnly' isolation criterion, which requires no additional high-p_{T} tracks in a cone around the muon track. To ensure that the probe muon lies in the trigger efficiency plateau, it must have a p_{T} at least 5% above the p_{T} threshold of the trigger, therefore leading muon must have $p_{\mathrm{T}} > 27.3$ GeV. The subleading muon is required to have $p_{\mathrm{T}} > 10$ GeV. When filling tau-related histograms, the tau RNN is required to be larger than 0.02 for the tau in question.

5.2 The FF measurements in $Z(\mu\mu)$ + jets events

As already mentioned, the fake-τ candidates in $Z(\mu\mu)$ + jets samples come mostly from misidentifying quark-initiated jets. In Fig. 5.1 are displayed fractions of leading τ object truth matched to the either quark- or gluon-initiated jet, but for lower p_{T} , the significant fraction of leading τ is unmatched. The unmatched jets correspond to the objects that failed to be matched to any truth-object. The vast majority of them are pile-up jets, which are not kept at the truth level in MC. These fractions were measured for antiMedium working point that corresponds to RNN score < 0.25 with 1 prong and RNN score < 0.40 with 3 prong (see Tab. 2.1) and for five different decay modes: 1 prong and no π^0 (1p0n), 1 prong and 1 π^0 (1p1n), 1 prong and more than one π^0 (1pXn), 3 prong and no π^0 (3p0n), and 3 prong and at least one π^0 (3pXn). These pictures confirm that the quark-initiated jets dominate in the $Z(\mu\mu)$ + jets events.

5.2.1 True-Tau Dilution in $Z(\mu\mu)$ + jets events

To obtain $N^{\mathrm{SR}}_{\mathrm{failID}}$, $N^{\mathrm{CR}}_{\mathrm{failID}}$ and $N^{\mathrm{CR}}_{\mathrm{passID}}$, we need to subtract the number of events that contain $\tau_{\mathrm{had\_vis}}$ truth-matched to true taus in MC. The MC also contains not only the reconstructed objects but also real objects born in pp collisions that produced some signal in the detector. The truth-matching to $\tau_{\mathrm{had\_vis}}$ is the procedure in which we are trying to match reco τs with real τs or some other truth object. If we are successful in matching reco τ with truth τ, then we will assume that the reco τ was produced based on the detector's response to the flight of real τ. If we do not succeed, we will expect the reco τ to be fake. The requirement for truth matching is $\Delta R < 0.2$ between the truth object and reco $\tau_{\mathrm{had\_vis}}$, and if there are multiple objects within this cone, the object with the highest p_{T} is chosen.

Before subtracting the contributions from true leptons, we studied two approaches for this subtraction. In the first approach, the reco τs are matched to truth-τ, and in the second one, the reco τs are also matched to truth-τ, but moreover, the reco μs are matched to truth-μs. The comparison of these two approaches is shown in Fig. 5.2 for 1-prong and Fig. 5.3 for 3-prong . The histograms in both figures are plotted as a function of p_{T} for different $\tau_{\mathrm{had\_vis}}$ decay modes. Each of the pictures on the right-hand side corresponds to the Medium RNN working point and on the left-hand side to the antiMedium RNN working point. For all histograms,

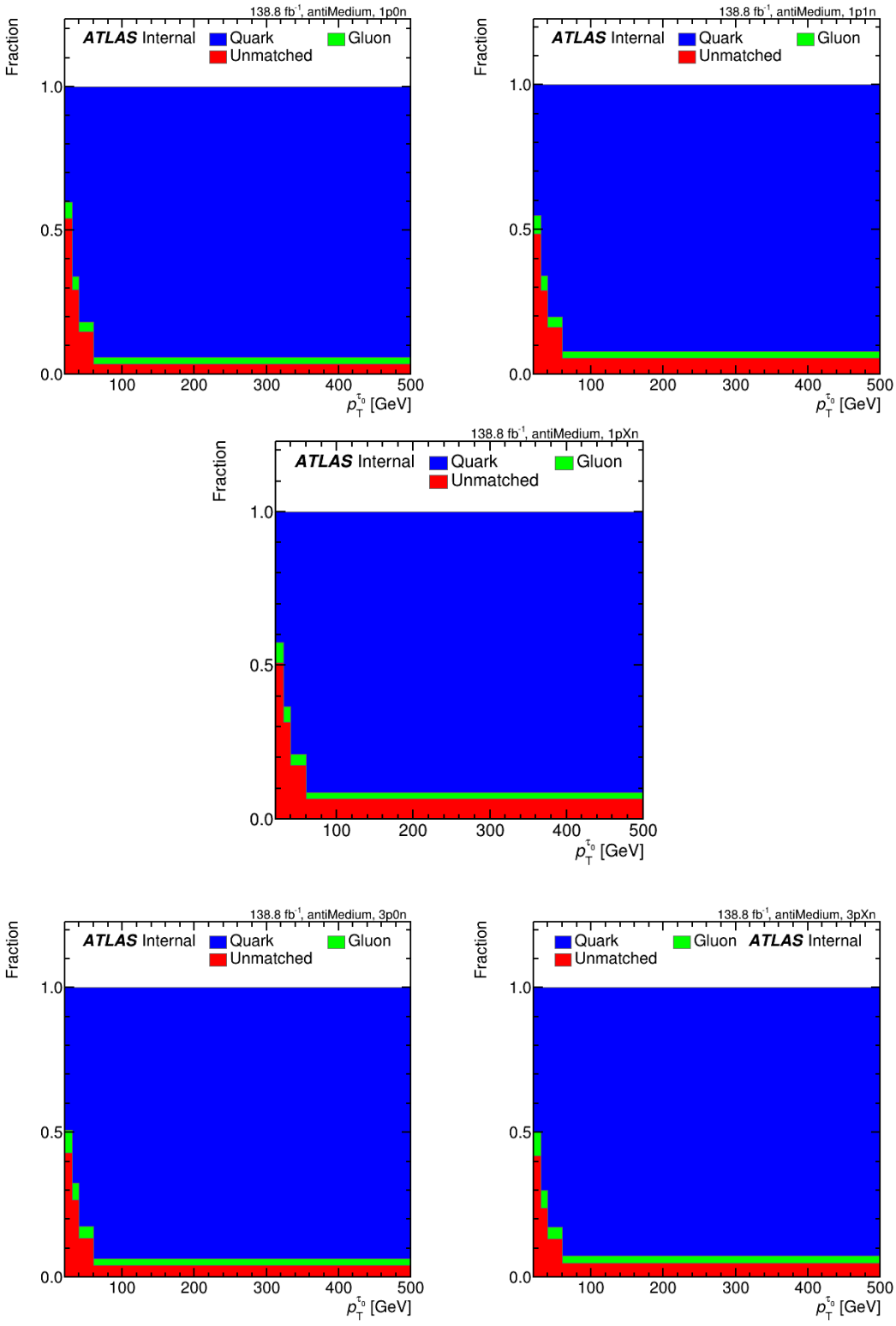


Figure 5.1: Fractions of leading τ object truth matched to the either quark- or gluon-initiated jet or unmatched in Z($\mu\mu$) + jets MC events plotted as function of leading τ p_{T} for different τ decay modes for antiMedium working point. The blue area corresponds to the quark fraction, the green one to the gluon fraction and red one to the unmatched leading τ objects

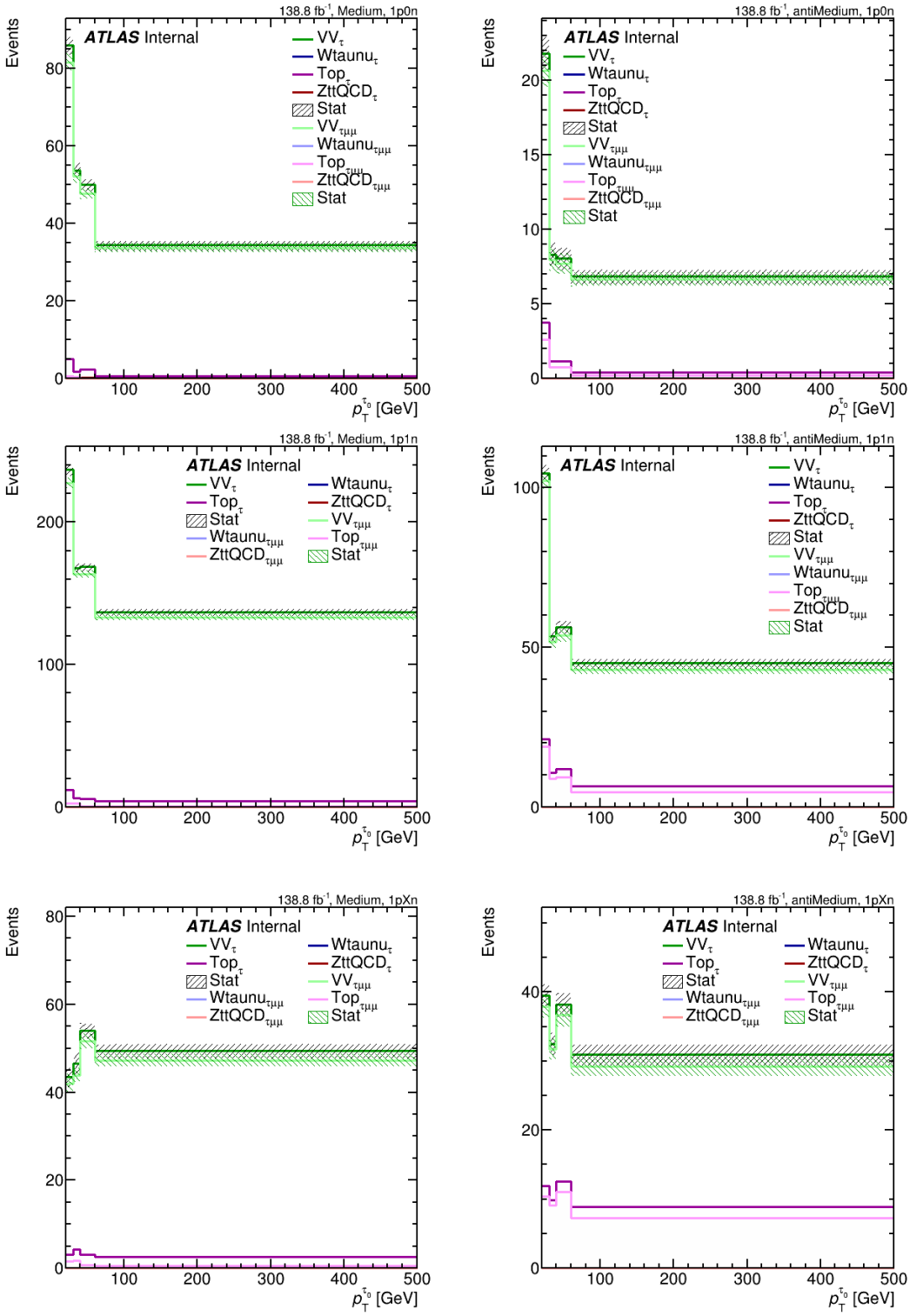


Figure 5.2: Comparison of two truth matching approaches: 1. reco τ matched to a truth τ, 2. reco τ matched to truth τ and reco μs matched to truth μ. Histograms are plotted as a function of p_{T} for one prong decay modes and Medium (left) and antiMedium (right) RNN ID working points. The contributions from MC samples are stack on each other. The lines with lightest colors corresponds to the approach with also truth-matching the μs. The hatched bands correspond to the statistical uncertainties

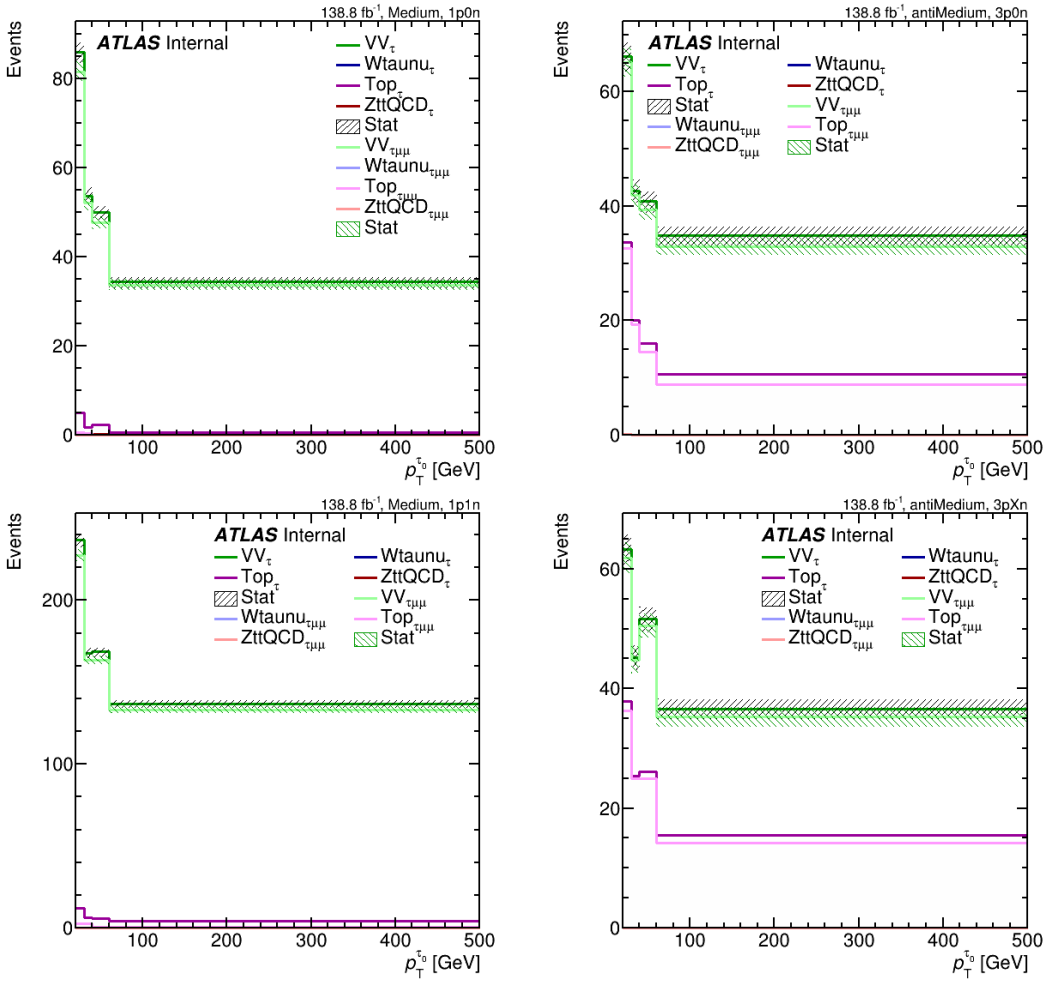


Figure 5.3: Comparison of two truth matching approaches: 1. reco τ matched to a truth τ, 2. reco τ matched to truth τ and reco μs matched to truth μ. Histograms plotted as a function of p_{T} for three prong decay modes and Medium (left) and antiMedium (right) RNN ID working points. The contributions from MC samples are stack on each other. The lines with lightest colors corresponds to the approach with also truth-matching the μs. The hatched bands correspond to the statistical uncertainties

it holds that the difference between truth-matching only τ and truth-matching also μs is negligible given the statistical uncertainty. With truth matching only τ and not μs, we also accept fake-μs, which is beneficial for us. Although MC is not able to simulate fakes correctly, it is better to have at least some estimate than none, and since the number of fake-μs is small, we can choose this approach. Therefore, for subtracting the contribution from true leptons, we decided to use the easier approaches with truth-matching the reco τ to truth-τ.

In figures 5.2 and 5.3, we can also notice that the major contribution for truth-matched τs comes from the diboson (VV) and Top MC samples.

With truth-matching only the reco τs to the truth-τs, true-τ dilution and data in $Z(\mu\mu)$ + fake $\tau_{\mathrm{had\_vis}}$ candidates samples are shown in Fig. 5.4 and Fig. 5.5 for both the Medium (histograms on the left-hand side) and antiMedium (histograms

on the right-hand side) RNN working points and each τ decay modes plotted as a function of p_{T} . In all cases, the true-τ dilution is low.

The low true-τ dilution is also responsible for the fact that the FFs that are measured only from data are within the statistical uncertainty equal to the FFs measured with the subtraction of contribution from true-τs. This is shown in Fig. 5.6, where the corresponding FFs with and without true-τ subtraction measured in $Z(\mu\mu)$ + jets events are plotted as a function of p_{T} for different τ decay modes and Medium RNN working point.

5.2.2 The FFs for Leading and Subleading Fake-τ Candidate

For FF measurement in $Z(\mu\mu)$ + jets events, we select at least one τ surviving the OLR and selection criteria. In Fig. 5.7 is shown the histogram for the number of τs in the event that survived selection criteria, and moreover, the blue histogram corresponds to the data without applying the cut on RNN score > 0.01, whilst the red histogram accounts for this cut on RNN score. These histograms correspond to the Medium RNN working point, and each of them is plotted for different τ decay mode. It can be seen that the number of events with more than one τ is much smaller than the number of events with exactly one τ.

Considering events with at least one fake-τ candidate, we looked at the importance of $Z(\mu\mu)$ +jets events with more than one fake-τ candidate. We measured the FF for the leading τ candidate caring only on properties of this candidate and FF for the subleading τ candidate and again taking into account only the properties of this candidate. These FFs are displayed in Fig. 5.8, where the FFs are plotted as a function of p_{T} for Medium RNN working point, and different τ decay modes. In all plots, it is visible that for the lower p_{T} bins, the FFs for subleading τ are lower than those for leading τ and have much stronger statistical uncertainty. The difference in the behavior of the FFs for leading and subleading τ may be caused by the greater fraction of the subleading τs originated from gluon-initiated jets than for the leading τ (see Fig. 5.1). The fraction of quark-, gluon-initiated, or unmatched jets for the subleading fake-τ candidate is shown in Fig. 5.9. The fractions are plotted as a function of p_{T} of subleading τ for different τ decay modes and Medium RNN working point. Similarly, as for the leading τ candidate, for lower p_{T} bins, the most significant fraction corresponds to the unmatched jets, but for the higher p_{T} bins, even when there is a higher number of subleading fake-τs originating from gluon-initiated jets or unmatched, the subleading fake-τ candidate originates mostly from a quark-initiated jet.

Since the sample for subleading τ candidates is more contaminated with gluon-initiated and unmatched jets, we chose not to use it for the FF measurement. Moreover, this sample is much smaller compared to the sample for leading τ candidates, and thus it does not have much impact on the accuracy of FF measurement. This can be seen from the fact that the FFs for subleading τ candidates have greater statistical uncertainty, and the combined FFs for leading and subleading τ are almost identical.

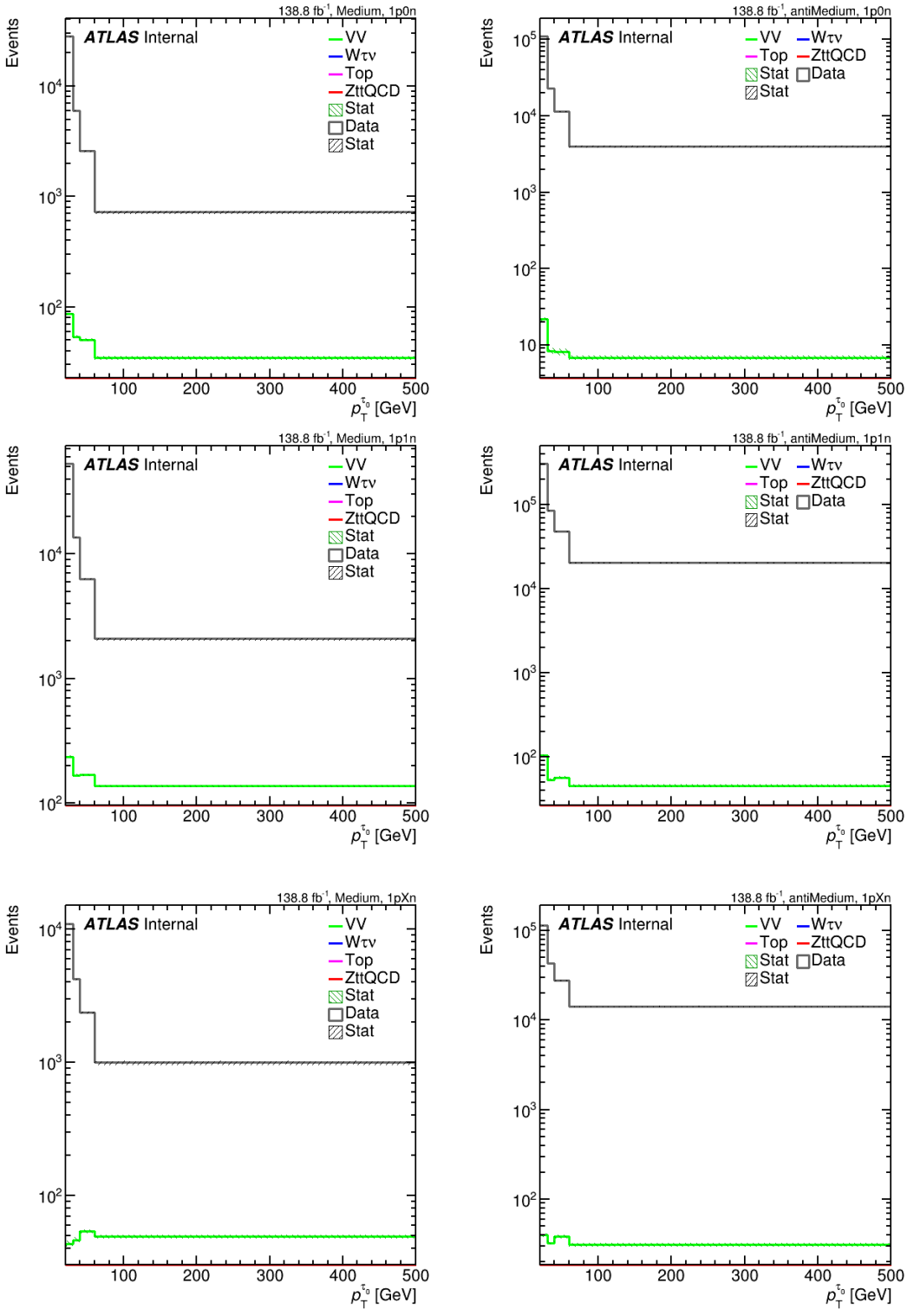


Figure 5.4: True Tau Dilution in $Z(\mu\mu)$ + $\tau_{\text{had\_vis}}$ candidate for different decay modes with one charged particle in Medium (left) and antiMedium (right) RNN ID region as a function of p_{T} . The black line corresponds to the data and the green one to the Diboson MC sample stack on top of the Top MC sample (magenta line). The hatched bands correspond to the statistical uncertainties.

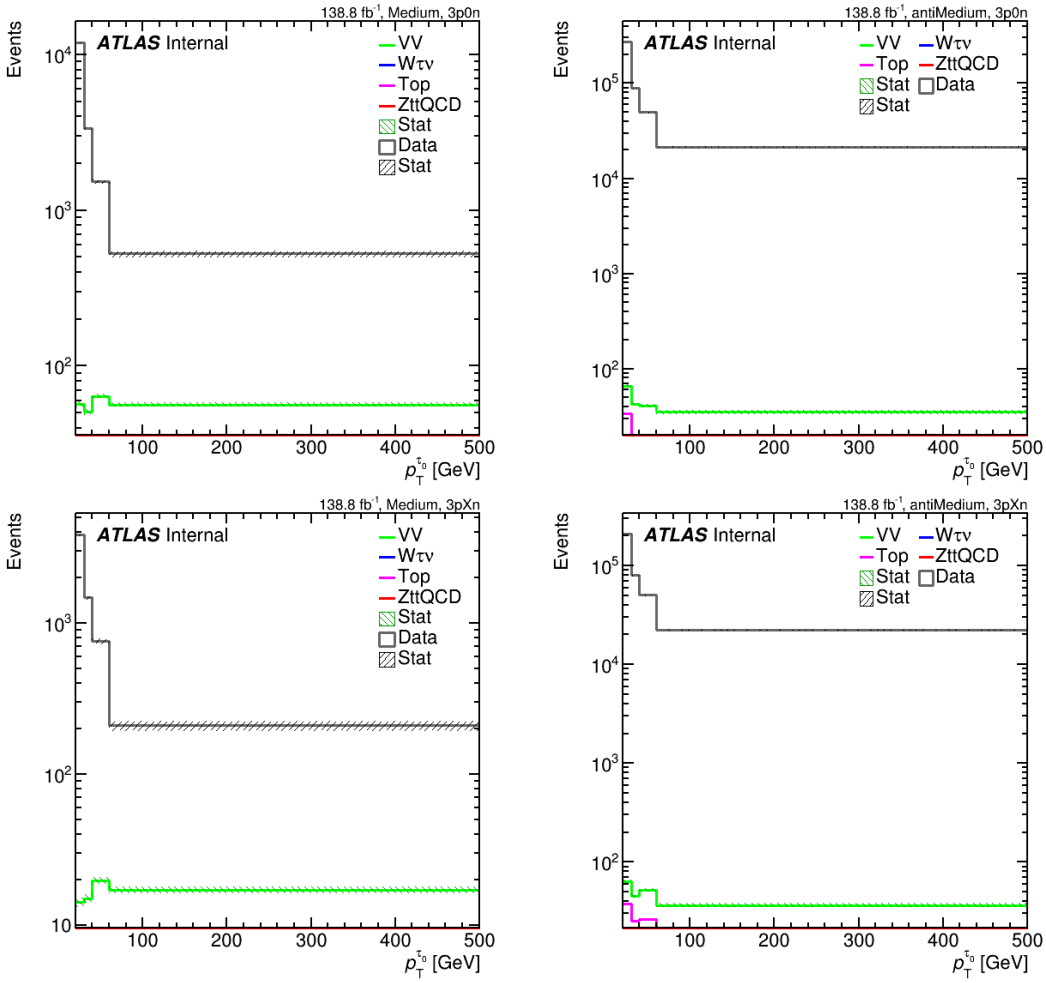


Figure 5.5: True Tau Dilution in $Z(\mu\mu)$ + $\tau_{\text{had\_vis}}$ candidate for different decay modes with three charged particles in Medium (left) and antiMedium (right) RNN ID region as a function of p_{T} . The black line corresponds to the data and the green one to the Diboson MC sample stack on top of the Top MC sample (magenta line). The hatched bands correspond to the statistical uncertainties.

5.2.3 Close-by Objects

In the case there is another object near the leading $\tau_{\text{had\_vis}}$ candidate, the leading candidate can be affected by this close-by object. To determine whether it is necessary to study the effects of close-by objects, we looked at the distance between the leading τ candidate and the nearest object among jets and τs in the $\eta \times \phi$ plane. In Fig. 5.10, the histograms for this study are shown, where we applied the cut on RNN score > 0.01 on all τ candidates (Fig. 5.10a) or only on the leading τ candidate (Fig. 5.10b). Since the jets were reconstructed using anti-k_{T} algorithm with radius parameter $R = 0.4$, and moreover, the hadronic τs have 'core' of $R = 0.2$, for both histograms holds that there is a negligible number of close-by objects to leading τ candidate, and hence, we will not deal with the effects caused by close-by jets or τs.

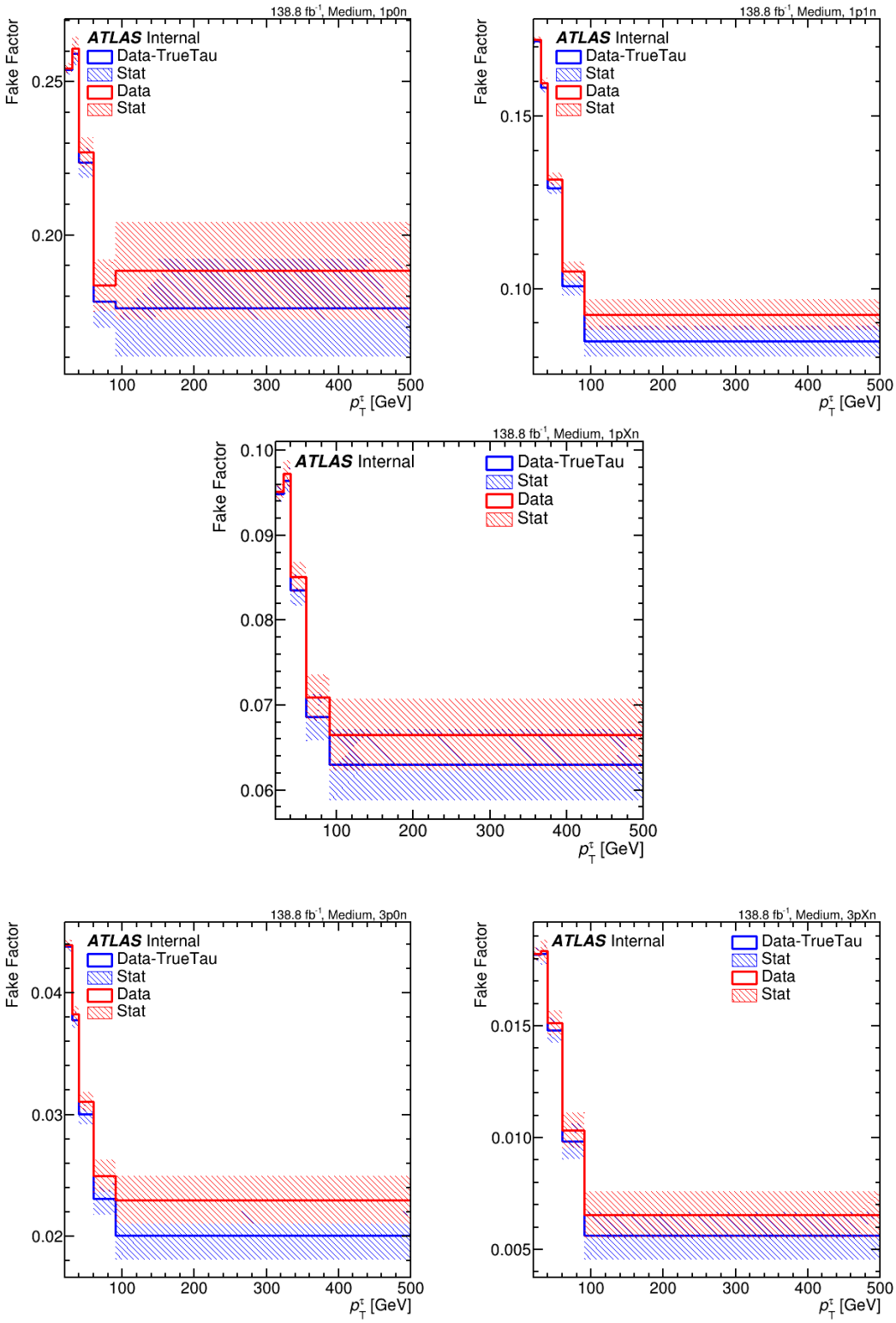


Figure 5.6: Comparison of FFs measured in $Z(\mu\mu)$+jets events with and without true τ subtraction for different τ decay modes for Medium ID working point. The FFs are plotted as a function of p_{T} and the blue one corresponds to the FF measured with true τ subtraction and the red one for FF without true τ subtraction. The hatched bands correspond to the statistical uncertainty

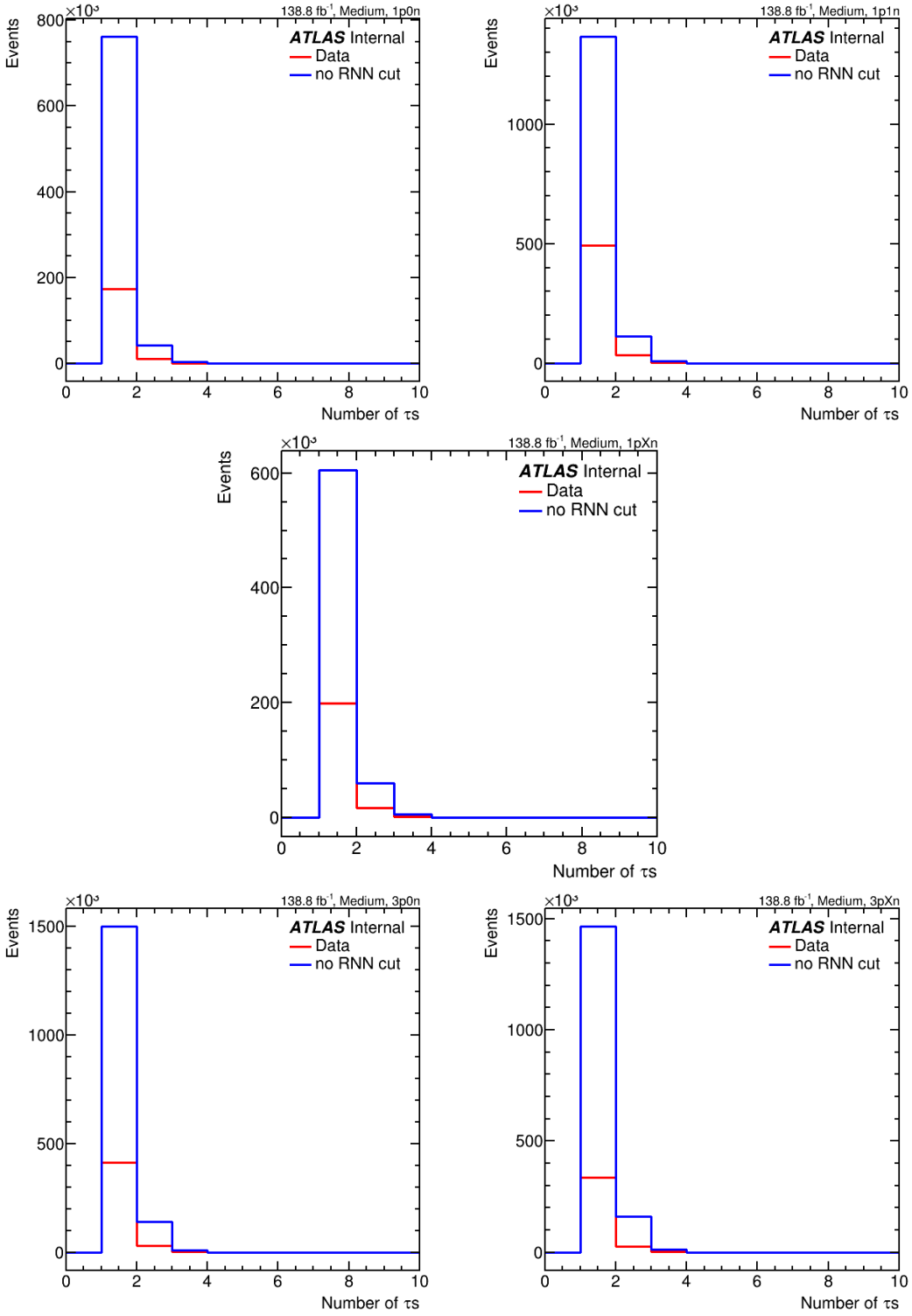


Figure 5.7: Number of fake τ candidates in $Z(\mu\mu)$+jets events with applying cut RNN$>$ 0.1 on leading τ candidate (red line) and without this cut (blue line)

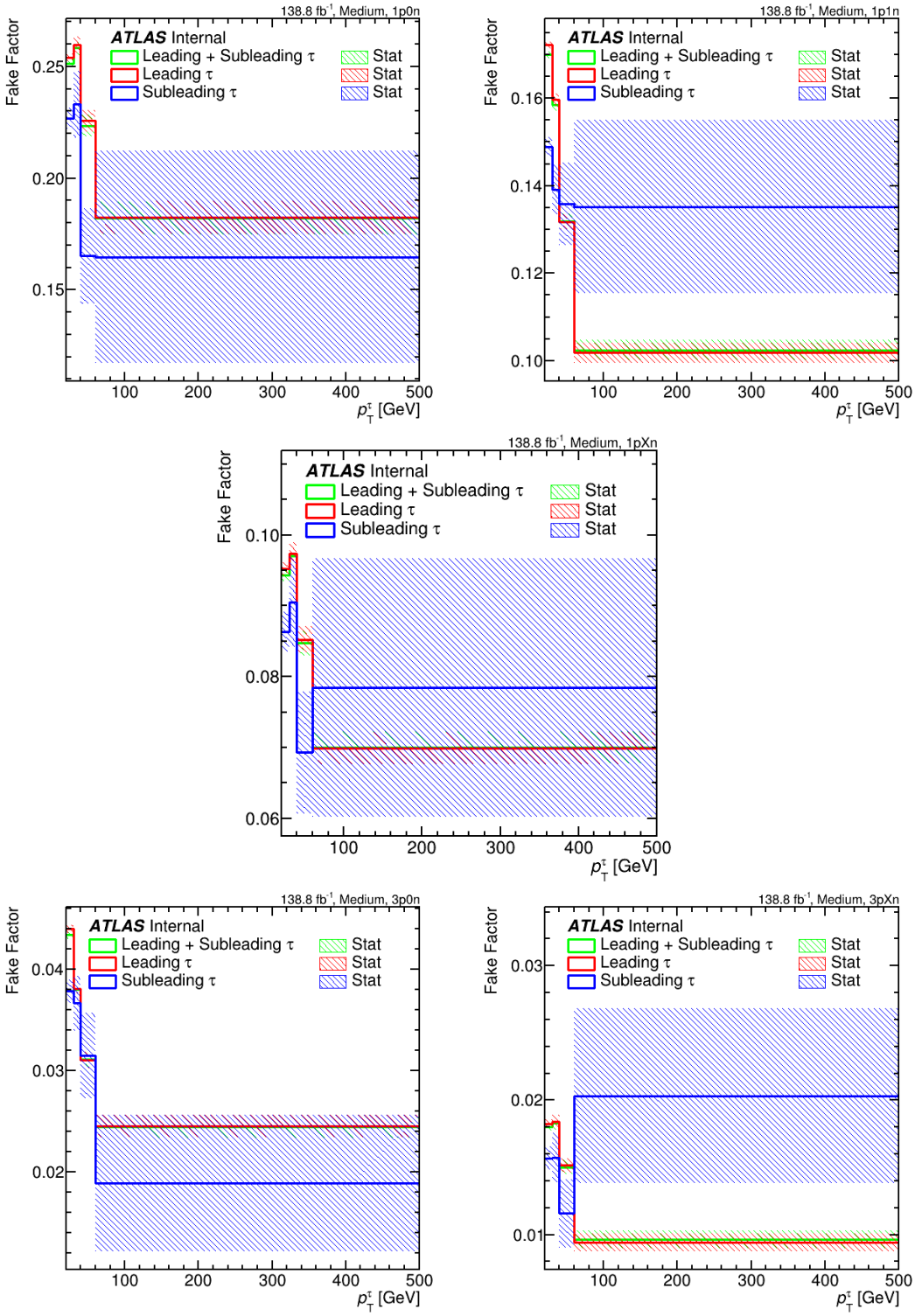


Figure 5.8: FFs measured in $Z(\mu\mu)$+jets events for leading fake τ candidate and subleading τ candidate for different decay modes and Medium ID RNN working point. The green line corresponds to the FF measured with the contribution from both leadning and sublieding $\tau_{\text{had\_vis}}$ candidates. The FF measured with contribution only from the leading $\tau_{\text{had\_vis}}$ candidates is plotted by red line and the blue line corresponds to the FF with contribution from subleading $\tau_{\text{had\_vis}}$ candidates. The hatched bands correspond to the statistical uncertainties.

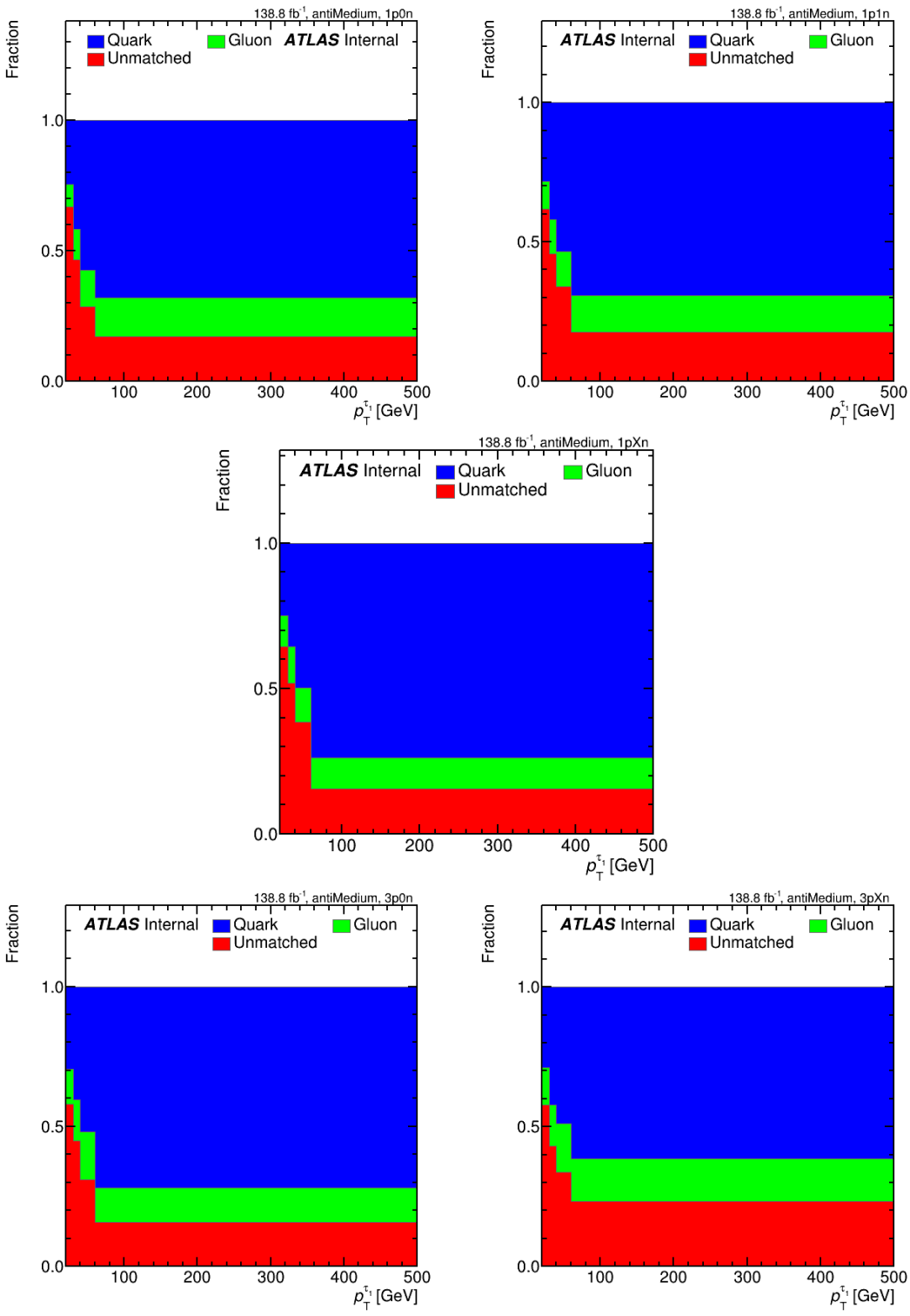
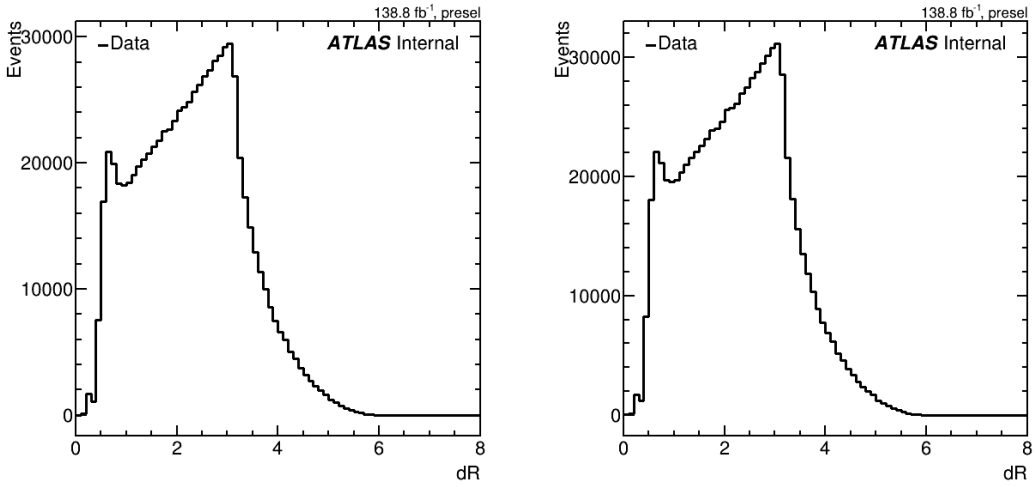


Figure 5.9: Fractions of subleading τ object truth matched to the either quark- or gluon-initiated jet or unmatched in $Z(\mu\mu)$ + jets MC events plotted as function of subleading τ p_T for different τ decay modes for antiMedium working point. The blue area corresponds to the quark fraction, the green one to the gluon fraction and red one to the unmatched subleading τ objects



(a) with RNN> 0.01 cut on all $\tau_{\text{had\_vis}}$ candidates

(b) with RNN> 0.01 cut on leading $\tau_{\text{had\_vis}}$ candidate only

Figure 5.10: Distance between leading τ candidate and the nearest object from jets and τs in $\eta \times \phi$ plane

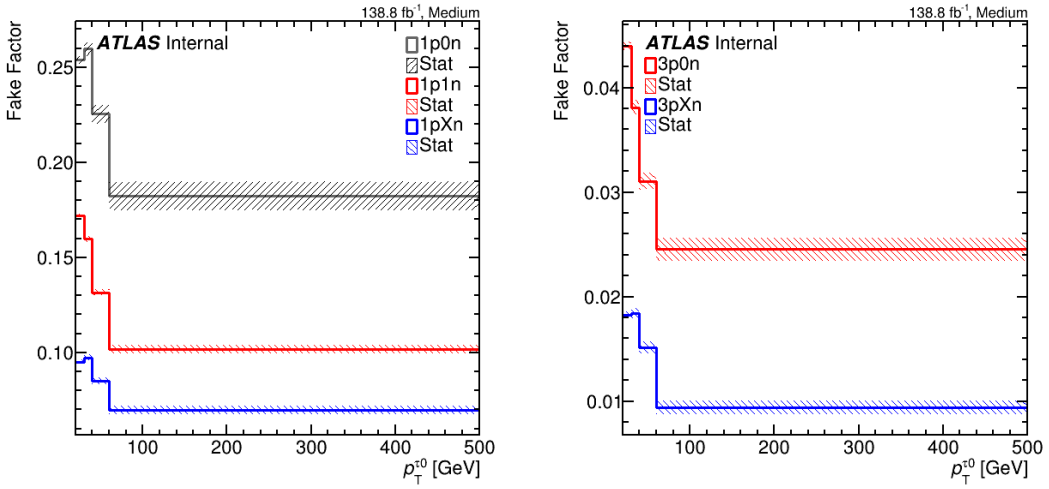


Figure 5.11: Dependence of FFs on τ decay modes for 1 prong (left) and 3 prong (right) as a function of p_{T} for Medium ID RNN working point. On the left side, the black line corresponds to the decay mode 1p0n, the red one is for decay mode 1p1n and the blue one corresponds to the decay mode 1pXn. On the right side, the red line corresponds to the $\tau_{\text{had\_vis}}$ decaying into 3p0n and the blue line is for decay mode 3pXn. The hatched bands correspond to the statistical uncertainties

5.2.4 Dependence of the Fake Factor

The FF can be dependent on many variables, such as p_{T} , η, μ, decay mode, where μ is the expected number of pp interactions in one event (bunch-crossing). Therefore for the FF measurement, we need to split each region into subregions, e.g., having $p_{\mathrm{T}} \in (20, 30)$ GeV, decay mode with 1 charged particle and 0 neutral π^0, and so on, and measure the FFs for each of these subregions. In this section, we will study the dependence of FF on these variables, according to which we determine the subregions.

A careful reader may have already noticed that all previous FFs have been plotted as a function of p_{T} for different τ decay modes and that the FF is truly strongly dependent on these two variables. For a less careful reader, this fact is shown in Fig. 5.11. In these pictures, the FFs are again plotted as a function of p_{T} , and each of the FF corresponds to a different decay mode with 1-prong on the left side and 3-prong on the right side.

We also studied the dependence of FF on η and μ. This is shown in Fig. 5.12, where the FFs are plotted as a function of p_{T} for different decay modes in the Medium ID region. Each of the plotted FFs corresponds to the different $\eta \times \mu$ subbin. There is a certain dependence on η primarily for lower p_{T} bins, which can be seen by the gray and blue slice, corresponding to $\eta \in (0, 1.37)$ being slightly far away from green and red one for $\eta \in (1.52, 2.5)$. The dependence on μ is weaker than for η.

5.2.5 Track-based jet width

In Chapter 4, we mentioned that the quark fraction for given regions needs to be determined using the template fit, and the suitable variable that will distinguish between quark- and gluon-initiated jets is track-based jet width. In Fig. 5.13-5.18, we studied if it is truly so.

In Fig.5.13, 5.14 and 5.15, there is illustrated the comparison between the track-based jet width shape of data and quark-, gluon-initiated and unmatched jets in Z($\mu\mu$)+jets MC samples for $p_{\mathrm{T}} \in (20, 30)$ GeV, $p_{\mathrm{T}} \in (30, 40)$ GeV and $p_{\mathrm{T}} \in (40, 60)$ GeV, respectively, in antiMedium ID region for different decay modes. These spectra are normalized to one. There is a visible difference between track-based jet width shape for quark- and gluon-initiated jets and the unmatched jets spectrum. The unmatched template has a stronger high-jet-width tail than the quark and gluon templates.

By the first look, we are not able to tell if the data shape looks similar to some other template. Therefore these MC templates would have to be fitted on data in order to draw some conclusions. However, with the increasing p_{T} , the shape of the track-based jet width spectrum for data is more similar to the shape of the quark-initiated template. This is expected because with the increasing p_{T} , the significance of the unmatched (pile-up) jets decreases (see Fig. 5.9).

The negative track width is attributed to the jets that do not contain any track

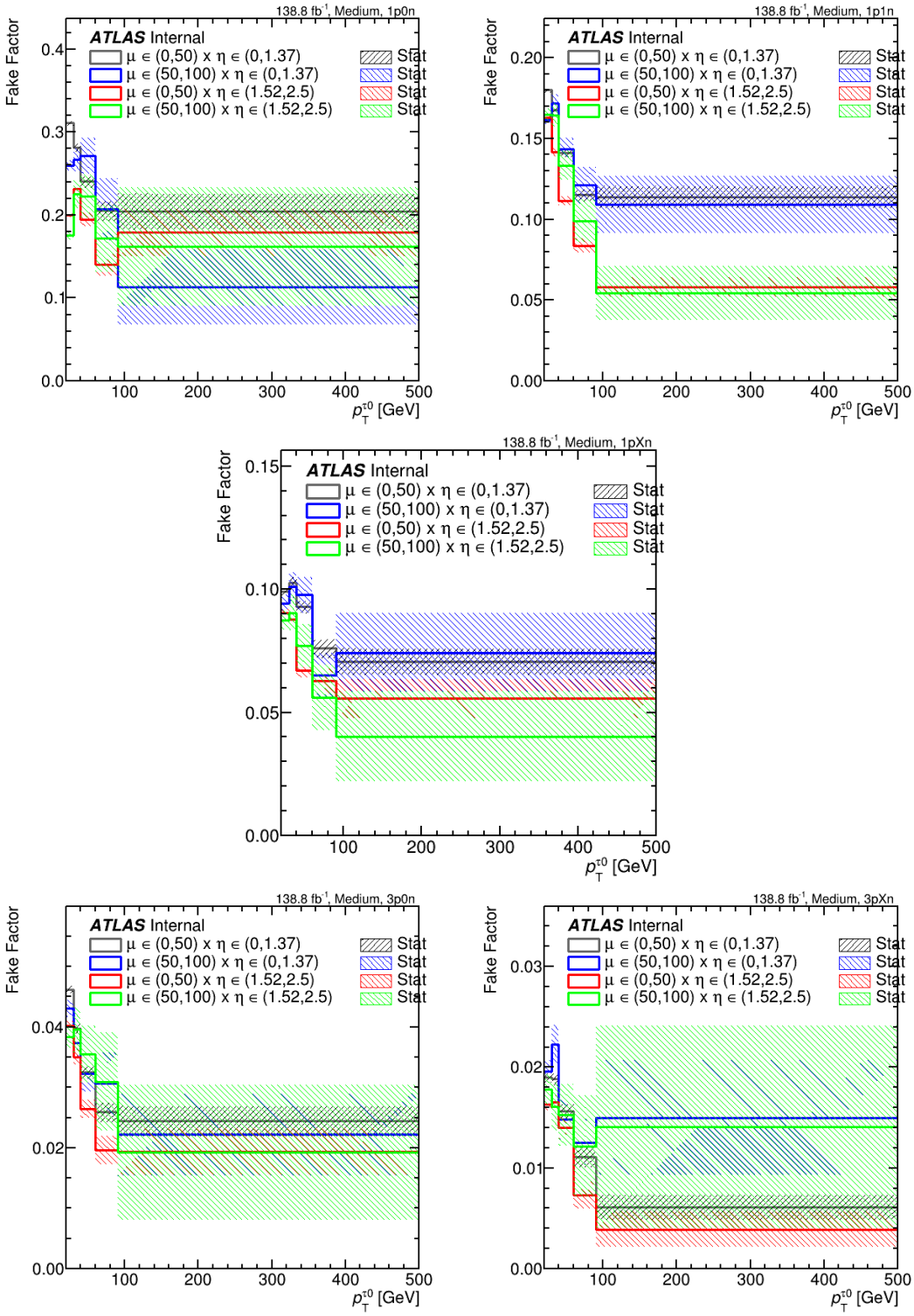


Figure 5.12: Studying the dependence of FF on μ and η. The FFs are plotted as a function of p_{T} for different τ decay modes for the Medium ID working point. Each of the slice corresponds to the different $\eta \times \mu$ subbin: the gray one for $\mu \in (0, 50) \times \eta \in (0, 1.37)$, the blue one for $\mu \in (50, 100) \times \eta \in (0, 1.37)$, the red one for $\mu \in (0, 50) \times \eta \in (1.52, 2.5)$ and green one for $\mu \in (50, 100) \times \eta \in (1.52, 2.5)$. The hatched bands correspond to the statistical uncertainties.

with $p_{\mathrm{T}} > 1$ GeV associated with tau vertex. In the negative bin, unmatched jets are completely dominant, which only underlines that they are pile-up jets (pile-up jet contains mainly tracks leading to the vertex, where the given jet was born, and therefore not to the tau vertex). The gluon-initiated jets are more represented in the negative jet width bin than quark-initiated jets, which is probably due to the fact that gluon-initiated jets contain, on average, more charged particles with lower p_{T}, while quark jets contain less charged particles, and they then have each larger p_{T}. Thus, in gluon-initiated jets, it is more likely that no track has p_{T} > 1 GeV.

Fig.5.16,5.17,5.18, show the fraction of quark-, gluon-initiated and unmatched jets n Z($\mu\mu$)+jets MC samples in bins of track-based jet width for antiMedium working point, different decay modes and for $p_{\mathrm{T}} \in (20, 30)$ GeV, $p_{\mathrm{T}} \in (30, 40)$ GeV and $p_{\mathrm{T}} \in (40, 60)$ GeV, respectively. These pictures are complementary to the comparison of the track-based jet width shape shown in Fig.5.13−5.15. They show, for each jet width bin, what the quark/gluon/unmatched jet fraction is in it. From this, we can see that in the bin with the negative jet width, there is always at least 70% of the unmatched jets. We also clearly see how the percentage of unmatched jets decreases in all bins with increasing p_{T} and also that the percentage of quark jets is always significantly higher than the percentage of gluon jets, which we, of course, expected from $Z(\mu\mu)$ + jet events.

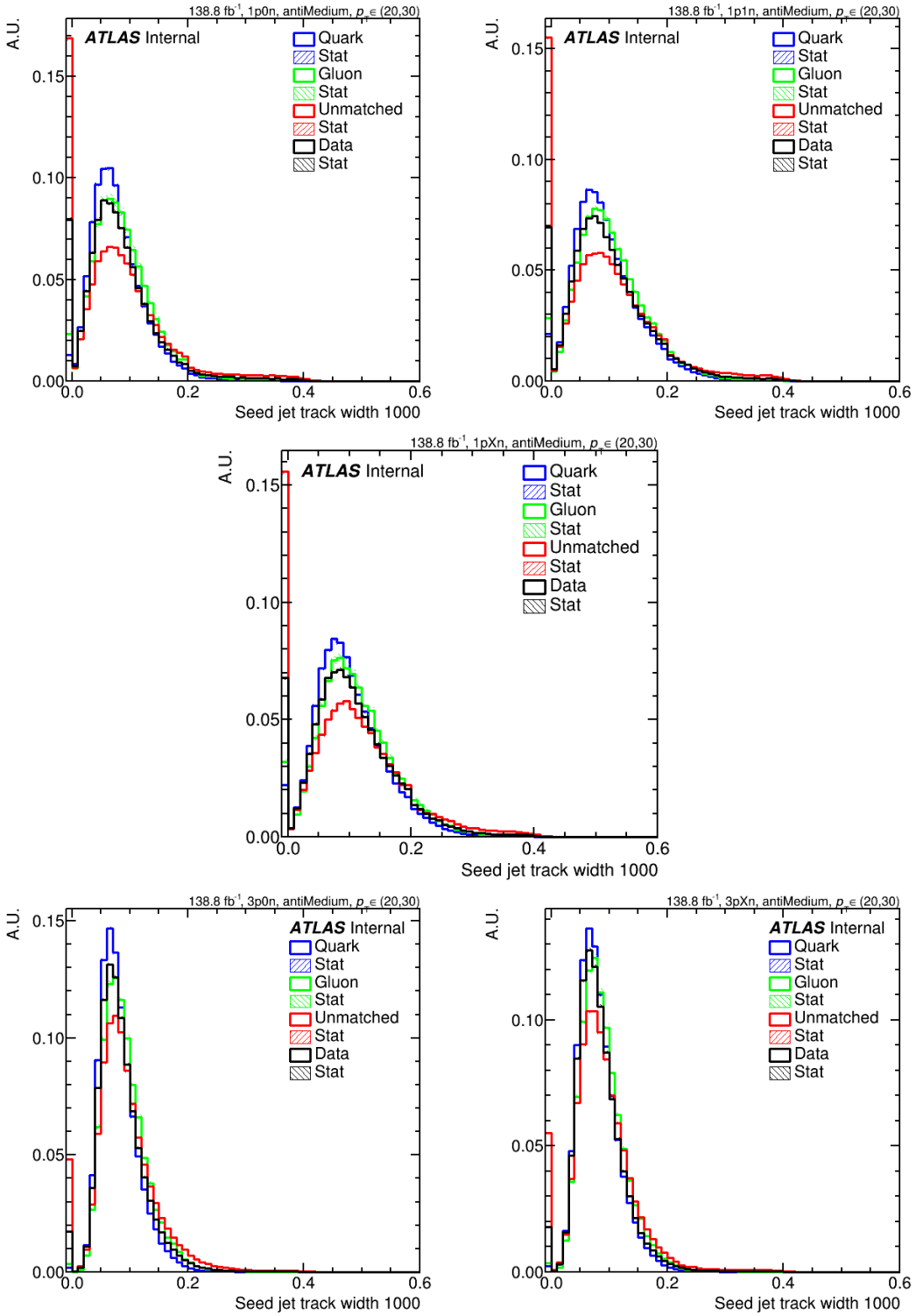


Figure 5.13: Track-based jet width shape comparison of data and quark-, gluon-initiated and unmatched jets in Z($\mu\mu$) + jets MC events for $p_{\mathrm{T}} \in (20, 30)$ for antiMedium ID working point. The black spectrum corresponds to the data, the blue one is for quark template, the green one for gluon template and the red one for unmatched. The hatched bands correspond to the statistical uncertainties

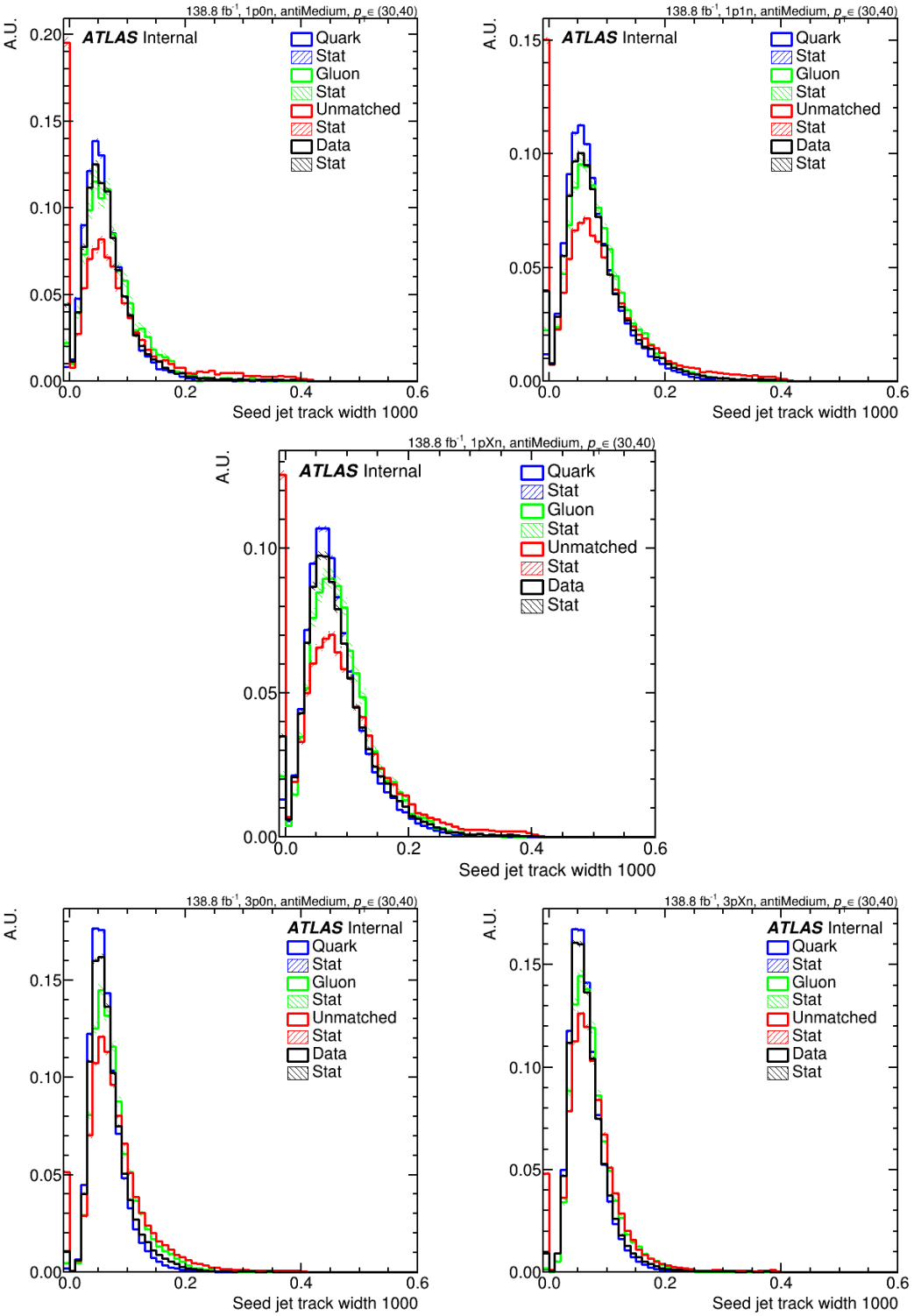


Figure 5.14: Track-based jet width shape comparison of data and quark-, gluon-initiated and unmatched jets in Z($\mu\mu$) + jets MC events for $p_{\mathrm{T}} \in (30, 40)$ for antiMedium ID working point. The black spectrum corresponds to the data, the blue one is for quark template, the green one for gluon template and the red one for unmatched. The hatched bands correspond to the statistical uncertainties

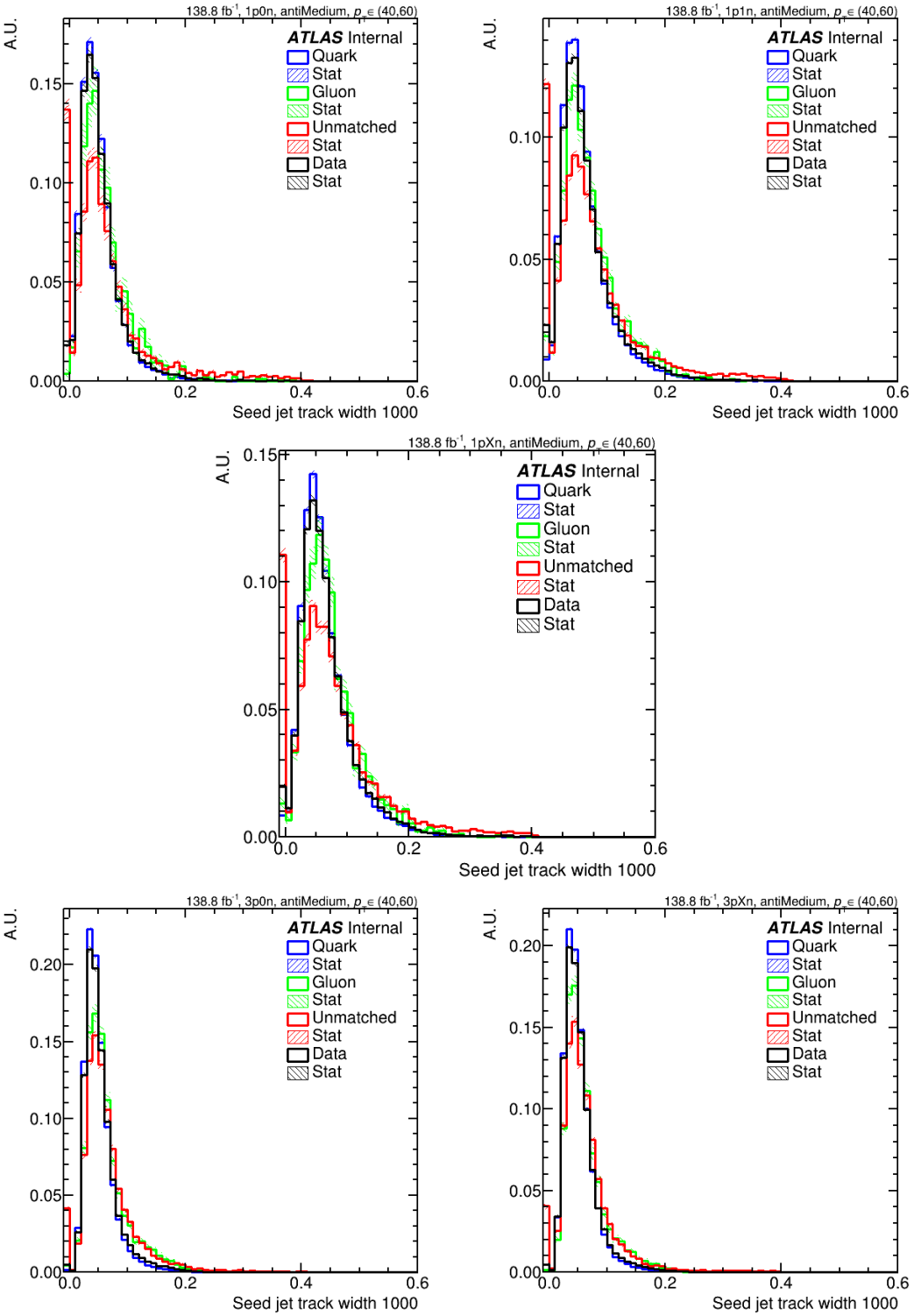


Figure 5.15: Track-based jet width shape comparison of data and quark-, gluon-initiated and unmatched jets in Z($\mu\mu$) + jets MC events for $p_{\mathrm{T}} \in (40, 60)$ for antiMedium ID working point. The black spectrum corresponds to the data, the blue one is for quark template, the green one for gluon template and the red one for unmatched. The hatched bands correspond to the statistical uncertainties

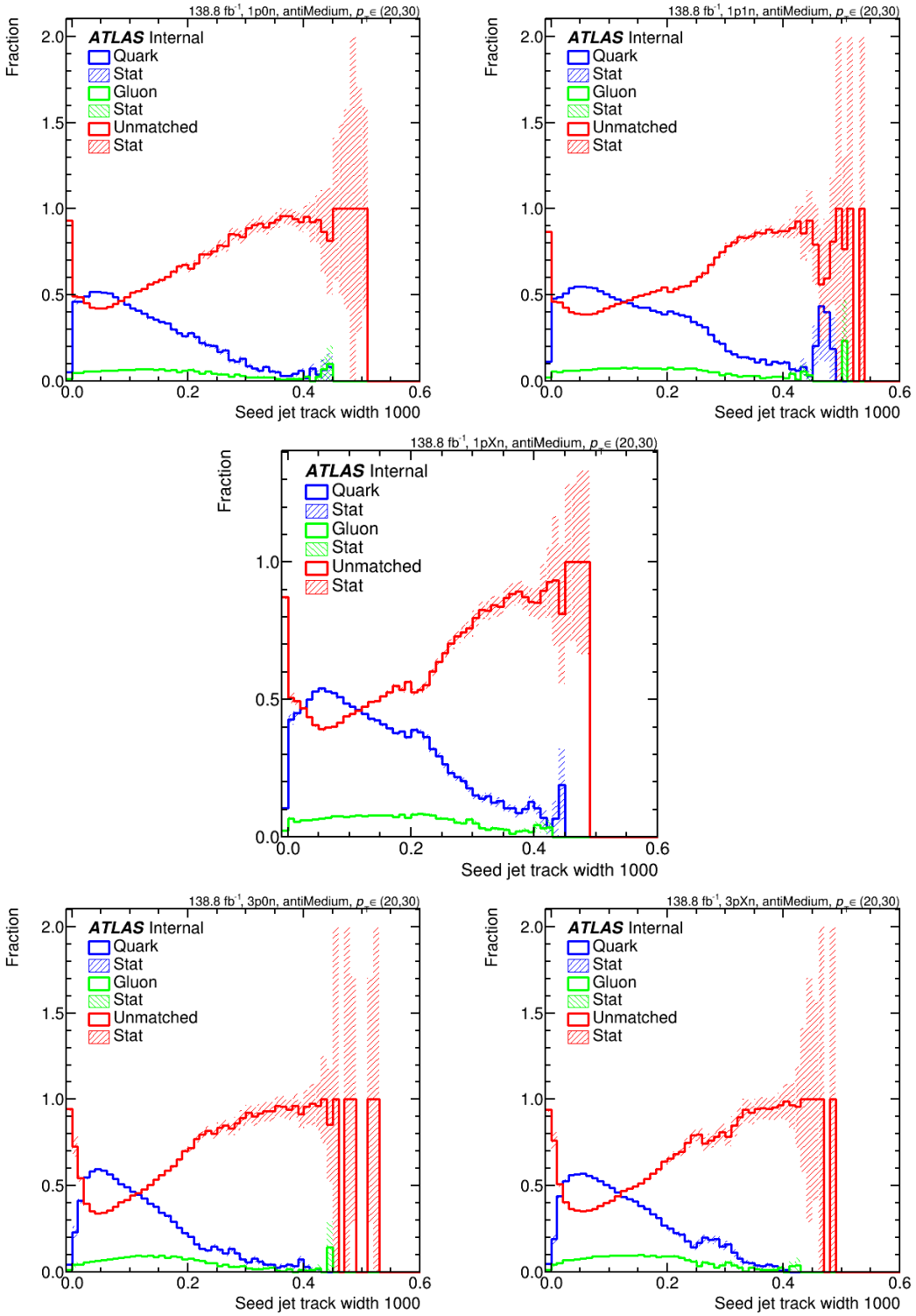


Figure 5.16: Fraction of quark-, gluon-initiated and unmatched jets in bins of track-based jet width in $Z(\mu\mu)$ + jets MC events for $p_T \in (20, 30)$ and antiMedium ID working point. The blue spectrum corresponds to quark-initiated jets, the green one is for gluon-initiated jets and the red one for unmatched. The hatched bands correspond to the statistical uncertainties

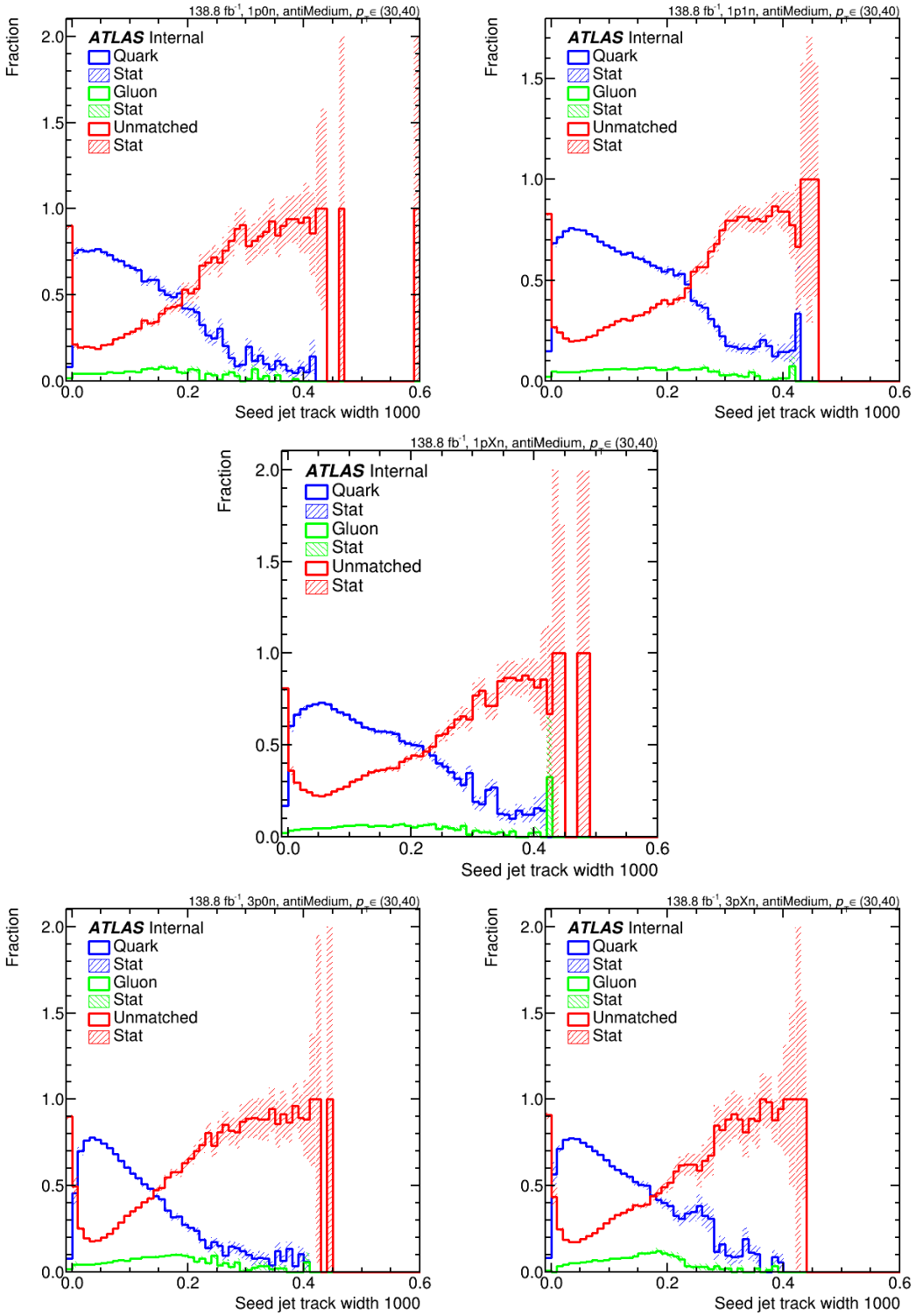


Figure 5.17: Fraction of quark-, gluon-initiated and unmatched jets in bins of track-based jet width in $Z(\mu\mu)$ + jets MC events for $p_T \in (30, 40)$ and antiMedium ID working point. The blue spectrum corresponds to quark-initiated jets, the green one is for gluon-initiated jets and the red one for unmatched. The hatched bands correspond to the statistical uncertainties

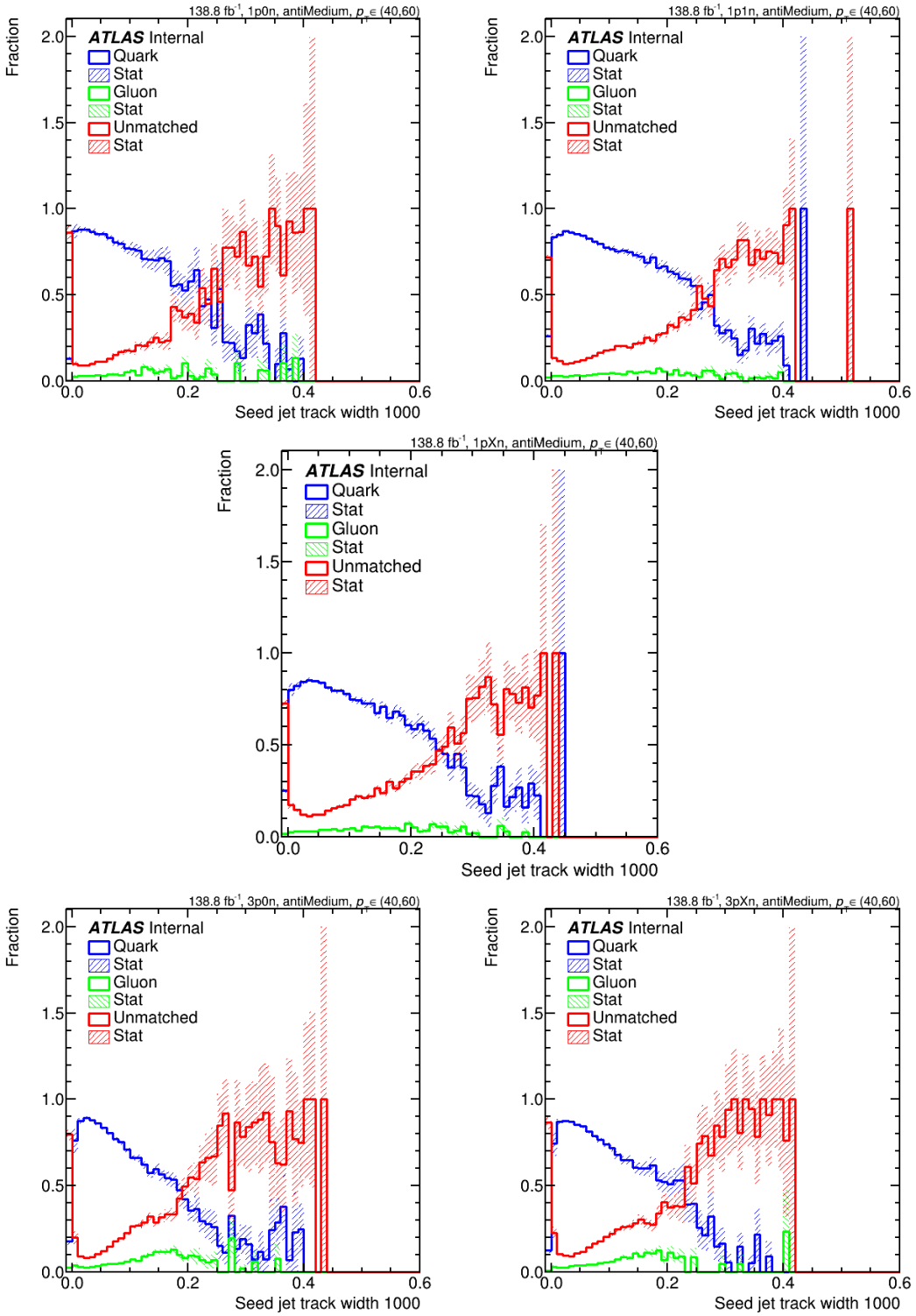


Figure 5.18: Fraction of quark-, gluon-initiated and unmatched jets in bins of track-based jet width in Z($\mu\mu$) + jets MC events for $p_{\mathrm{T}} \in (40, 60)$ and antiMedium ID working point. The blue spectrum corresponds to quark-initiated jets, the green one is for gluon-initiated jets and the red one for unmatched. The hatched bands correspond to the statistical uncertainties

Conclusion

Within this thesis, the fake factor method for estimating fake-τ background was introduced. We also introduced the universal fake factor method that accounts for the quark/gluon composition in any particular SR.

We measured the FFs in $Z(\mu\mu)+$ jets events. The $\tau_{\text{had\_vis}}$ candidates in these events originate mainly from quark-initiated jets. However, there is a higher gluon fraction for subleading $\tau_{\text{had\_vis}}$ candidate. In lower p_{T} bins, the fraction for unmatched jets is more significant. The true-tau dilution is low. We studied two different approaches for truth-matching, and for further measurements, we decided to match only the reco τ to truth τ. We also measured the FFs for subleading τ candidate. These FFs are smaller than the one for the leading τ candidate, have larger statistical uncertainty, plus are more contaminated by gluon-initiated jets. Therefore, for FF measurements, we prefer working only with the leading τ candidate. The number of close-by objects in Z+jets events is negligibly small, and thus we do not consider the effects of close-by objects. The study of track-based jet width shape for quark-, gluon-initiated and unmatched jets shows the difference between these templates. The track-based jet width shape for data is becoming more similar to the quark template with increasing p_{T} , but still, the MC templates would have to be fitted on data in order to draw some conclusions. Measured FFs have a strong dependence on p_{T} and $\tau_{\text{had\_vis}}$ decay modes. There is also certain dependence on η and weaker dependence on μ, except for the lowest p_{T} bins.

Bibliography

[1] J. Horejsi and U. Karlova, *Fundamentals of Electroweak Theory*. Karolinum Press, 2002. https://books.google.cz/books?id=JGkFAQAACAAJ.

[2] J. Chýla, *Quarks, partons and Quantum Chromodynamics*, https://www.fzu.cz/~chyla/lectures/text.pdf.

[3] *Standard Model of particle physics*, https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg#file.

[4] Particle Data Group Collaboration, P. Zyla et al., *Review of Particle Physics*, PTEP **2020** (2020) no. 8, 083C01.

[5] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv:0802.1189 [hep-ph].

[6] E. Mobs, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, . https://cds.cern.ch/record/2197559.

[7] LHC Guide, Mar, 2017. https://cds.cern.ch/record/2255762.

[8] *A vacuum as empty as interstellar space*, CERN Accelerating Science. https://home.cern/science/engineering/vacuum-empty-interstellar-space.

[9] *The Large Hadron Collider*, CERN Accelerating Science. https://home.cern/science/accelerators/large-hadron-collider.

[10] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008) no. 08, . https://doi.org/10.1088/1748-0221/3/08/s08003.

[11] G. Aad et al., *ATLAS pixel detector electronics and sensors*, JINST **3** (2008) P07007. https://cds.cern.ch/record/1119279.

[12] A. Vogel, *ATLAS Transition Radiation Tracker (TRT): Straw Tube Gaseous Detectors at High Rates*, ATL-INDET-PROC-2013-005, CERN, Geneva, Apr, 2013. https://cds.cern.ch/record/1537991.

[13] ATLAS Collaboration, *Operation of the ATLAS trigger system in Run 2*, Journal of Instrumentation **15** (2020) no. 10, . http://dx.doi.org/10.1088/1748-0221/15/10/P10004.

[14] ATLAS Collaboration, *Luminosity Public Results Run2*, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2#Pileup_Interactions_and_Data_Tak.

[15] ATLAS Collaboration, M. zur Nedden, *The Run-2 ATLAS Trigger System: Design, Performance and Plan*, ATL-DAQ-PROC-2016-039, CERN, Geneva, Dec, 2016. https://cds.cern.ch/record/2238679.

[16] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. **C77** (2017) 490, arXiv:1603.02934 [hep-ex].

[17] T. Barillari et al., *Local Hadronic Calibration*, ATL-LARG-PUB-2009-001-2., CERN, Geneva, Jun, 2008. https://cds.cern.ch/record/1112035.

[18] ATLAS Collaboration, M. Aaboud et al., *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. C **77** (2017) no. 7, 466, arXiv:1703.10485 [hep-ex].

[19] ATLAS Collaboration, *Particle flow jet energy scale in 2015-2017 data and simulation*, Feb, 2019. https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2019-02/.

[20] ATLAS Collaboration, *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at* $\sqrt{s} = 13$ TeV, ATLAS-CONF-2017-029, CERN, Geneva, May, 2017. https://cds.cern.ch/record/2261772.

[21] ATLAS Collaboration, G. Aad et al., *Reconstruction of hadronic decay products of tau leptons with the ATLAS experiment*, Eur. Phys. J. C **76** (2016) no. 5, 295, arXiv:1512.05955 [hep-ex].

[22] ATLAS Collaboration, *Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment*, ATL-PHYS-PUB-2019-033, CERN, Geneva, Aug, 2019. https://cds.cern.ch/record/2688062.

[23] ATLAS Collaboration, G. Aad et al., *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at* $\sqrt{s}$ *=13 TeV*, Eur. Phys. J. C **76** (2016) no. 5, 292, arXiv:1603.05598 [hep-ex].

[24] ATLAS Collaboration, M. Aaboud et al., *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at* $\sqrt{s}$ *= 13 TeV*, Eur. Phys. J. C **79** (2019) no. 8, 639, arXiv:1902.04655 [physics.ins-det].

[25] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, , Jun, 2016. https://cds.cern.ch/record/2157687.

[26] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, arXiv:1002.2581 [hep-ph].

[27] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv:hep-ph/0409146.

[28] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv:0709.2092 [hep-ph].

[29] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C. P. Yuan, *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) 074024, arXiv:1007.2241 [hep-ph].

[30] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A Precision tool for QED corrections in Z and W decays*, Eur. Phys. J. C **45** (2006) 97–107, arXiv:hep-ph/0506026.

[31] S. Frixione, P. Nason, and G. Ridolfi, *A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*, JHEP **09** (2007) 126, arXiv:0707.3088 [hep-ph].

[32] E. Re, *Single-top Wt-channel production matched with parton showers using the POWHEG method*, Eur. Phys. J. C **71** (2011) 1547, arXiv:1009.2450 [hep-ph].

[33] S. Alioli, P. Nason, C. Oleari, and E. Re, *NLO single-top production matched with shower in POWHEG: s- and t-channel contributions*, JHEP **09** (2009) 111, arXiv:0907.4076 [hep-ph]. [Erratum: JHEP 02, 011 (2010)].

[34] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159–177, arXiv:1410.3012 [hep-ph].

[35] *ATLAS Pythia 8 tunes to 7 TeV data*, , CERN, Geneva, Nov, 2014. https://cds.cern.ch/record/1966419.

[36] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244–289, arXiv:1207.1303 [hep-ph].

[37] NNPDF Collaboration, R. D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, arXiv:1410.8849 [hep-ph].

[38] S. Frixione, E. Laenen, P. Motylinski, B. R. Webber, and C. D. White, *Single-top hadroproduction in association with a W boson*, JHEP **07** (2008) 029, arXiv:0805.3067 [hep-ph].

[39] Sherpa Collaboration, E. Bothmann et al., *Event Generation with Sherpa 2.2*, SciPost Phys. **7** (2019) no. 3, 034, arXiv:1905.09127 [hep-ph].

[40] T. Gleisberg and S. Hoeche, *Comix, a new matrix element generator*, JHEP **12** (2008) 039, arXiv:0808.3674 [hep-ph].

[41] F. Buccioni, J.-N. Lang, J. M. Lindert, P. Maierhöfer, S. Pozzorini, H. Zhang, and M. F. Zoller, *OpenLoops 2*, Eur. Phys. J. C **79** (2019) no. 10, 866, arXiv:1907.13071 [hep-ph].

[42] F. Cascioli, P. Maierhofer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012) 111601, arXiv:1111.5206 [hep-ph].

[43] A. Denner, S. Dittmaier, and L. Hofer, *Collier: a fortran-based Complex One-Loop LIbrary in Extended Regularizations*, Comput. Phys. Commun. **212** (2017) 220–238, arXiv:1604.06792 [hep-ph].

[44] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP **03** (2008) 038, arXiv:0709.1027 [hep-ph].

[45] S. Hoeche, F. Krauss, M. Schonherr, and F. Siegert, *A critical appraisal of NLO+PS matching methods*, JHEP **09** (2012) 049, arXiv:1111.1220 [hep-ph].

[46] S. Agostinelli et al., *Geant4—a simulation toolkit*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **506** (2003) no. 3, 250–303.

[47] ATLAS Collaboration, G. Aad et al., *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823–874, arXiv:1005.4568 [physics.ins-det].

[48] ATLAS Collaboration, M. Bahmani, *Data-driven estimation of fake tau background in Higgs searches in ATLAS*, , CERN, Geneva, Jan, 2019. https://cds.cern.ch/record/2653533.

[49] J. Alison, *The Road to Discovery*, ch. The Fake Factor Method, pp. 151–209. Springer International Publishing, 09, 2015.

[50] *Run2JetMoments*, 2019. https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/Run2JetMoments.

[51] ATLAS Collaboration, G. Aad et al., *Performance of the ATLAS muon triggers in Run 2*, JINST **15** (2020) no. 09, P09015, arXiv:2004.13447 [physics.ins-det].

List of Figures